UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA

QRadar UBA: Deteção e Análise de Anomalias Comportamentais de Segurança em Utilizadores

João Guilherme Cercas Filipe

Mestrado em Segurança Informática

Trabalho de projeto orientado por:
Prof. Doutor Pedro Miguel Frazão Fernandes Ferreira

2020

UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Informática

QRadar UBA: Deteção e Análise de Anomalias Comportamentais de Segurança em Utilizadores

João Guilherme Cercas Filipe

Trabalho de Projeto

Mestrado em Segurança Informática

Trabalho de projeto orientado por Prof. Doutor Pedro Miguel Frazão Fernandes Ferreira e supervisionado na instituição Altice Portugal por Eng. José António dos Santos Alegria.

2020

Agradecimentos

Agradeço ao Professor Doutor Pedro Ferreira por orientar este projeto, pelos preciosos conselhos e observações, e por se mostrar sempre disponível para ajudar e contribuir para o sucesso do mesmo.

Ao Engenheiro José Alegria, pela coorientação na Altice Portugal, agradeço a disponibilidade e confiança prestadas de forma incondicional, que contribuíram de forma decisiva para o sucesso deste trabalho.

A todos os elementos da Direção de Cyber Security e Privacidade da Altice Portugal, agradeço pela forma como fui acolhido e pelo seu essencial apoio na execução da fase de testes. Em particular, agradeço à equipa do SOC, nomeadamente ao Engenheiro Alberto Bruno, por ter estado sempre presente desde o planeamento até à concretização deste trabalho, pela sua colaboração e disponibilidade. Agradeço também aos elementos da Engenharia de Cibersegurança, Engenheiro Jorge Silva e Rui Almeida, por todo o apoio e disponibilidade demonstrados.

Por fim, agradeço à minha família, em especial à minha esposa Teresa, que através do seu apoio, paciência, capacidade de incentivo, disponibilidade e amor, formaram o suporte necessário para conseguir alcançar os objetivos propostos.

Para a minha esposa, Teresa, e para a minha filha, Carolina.

Resumo

Presentemente, vivemos num ambiente cada vez mais digital, onde as ameaças e ataques informáticos são uma constante. Com a crescente sofisticação dos ataques, e capacidade de reinvenção dos atacantes, emerge a necessidade de exploração de novos mecanismos e técnicas que visem uma deteção mais precoce e precisa. Mais ainda quando até os utilizadores que consideramos confiáveis, se podem traduzir numa ameaça interna. Neste contexto, a existência de ferramentas de *User and Entity Behavior Analytics* (UEBA) vem procurar responder a esta necessidade. Estas permitem construir um perfil de comportamento base por entidade e, assim, detetar qualquer desvio das suas atuais atividades, atribuindo um dado valor de risco.

Este projeto tem como objetivo provar a capacidade de deteção e análise de anomalias comportamentais dos utilizadores de uma grande organização, com impacto na cibersegurança dessa mesma organização, em tempo quase real. Para tal, far-se-á uso da plataforma *IBM QRadar* utilizada pelo *Cyber Security Operations Center* (CSOC) da Altice Portugal e, mais especificamente, a aplicação UBA, recorrendo às funcionalidades de *Machine Learning* por esta disponibilizadas.

Para alcançar o objetivo proposto, foi definida, primariamente, uma lista de casos de uso que endereçam as principais preocupações ao nível das possíveis ameaças internas, nomeadamente no que se refere ao possível abuso de credenciais e à existência de discrepâncias tempo-espaciais. Posteriormente, foram validadas quais as fontes que contribuem com os dados necessários para a monitorização dos casos de uso anteriormente identificados.

Este projeto torna possível produzir uma análise técnica crítica à aplicação UBA da plataforma *IBM QRadar*, no que respeita à deteção e análise de casos de uso concretos relevantes para a gestão da cibersegurança da Altice Portugal, nas áreas dependentes do comportamento dos seus utilizadores. Para mais, permite também avaliar o desempenho desta aplicação, sem comprometer o atual normal funcionamento da referida plataforma *IBM QRadar*.

Palavras-chave: Ameaça Interna, Comportamento, Deteção de Anomalias, UEBA, IBM QRadar, *Machine Learning*

Abstract

Nowadays, we live in an increasingly digital environment, where the threats and cyber attacks are always present. Considering the increasing attack sophistication and attacker reinvention capability, there is a need to explore new mechanisms and techniques for an earlier and more accurate detection, specially when users we find reliable can become an insider threat. In this context, tools like User and Entity Behavior Analytics (UEBA) come to give an answer to this concern. These tools allow you to build a baselining behavior profile by entity and, thus, any deviation from your current activities can be detected and a given risk value is assigned to that deviation.

This project aims to prove the ability to detect and analyze behavioral anomalies of users of a large organization, impacting in cyber security of that same organization, in near real time, making use of IBM QRadar used by the Cyber Security Operations Center (CSOC) of Altice Portugal, in particular the UBA App with its Machine Learning capabilities.

In order to achieve the proposed objective, a list of use cases has been defined, primarily, addressing the main concerns regarding possible insider threats, namely regarding the possible abuse of credentials and the existence of time-space discrepancies. Subsequently, the data sources that provide the essential data for monitoring the previously identified use cases were validated.

This project makes it possible to produce a critical technical analysis of the UBA App from IBM QRadar platform, as regards the detection and analysis of concrete use cases relevant to Altice Portugal's cybersecurity management, in areas dependent on the behavior of its users. Moreover, it also allows the evaluation of the performance of this application, in a way that it does not compromise the current normal operation of the IBM QRadar platform.

Keywords: Insider Threat, Behavior, Anomaly Detection, UEBA, IBM QRadar, Machine Learning

Conteúdo

Lista de Figuras

Lista de Tabelas

Abreviaturas

AD Active Directory.

API Application Programming Interface.

APT Advanced Persistent Threat.

AQL Ariel Query Language.

CISO Chief Information Security Officer.

CRE Custom Rules Engine.

CSIRT Computer Security Incident Response Team.

CSOC Cyber Security Operations Center.

CSV Comma-separated Values.

DCY Direção de Cyber Security e Privacidade.

DLP Data Loss Prevention.

DNS Domain Name Service.

DSM Device Support Module.

GPO Group Policy.

HTTP Hypertext Transfer Protocol.

ID Identificador.

IETF Internet Engineering Task Force.

IoC Indicators of Compromise.

IP Internet Protocol.

IPS Intrusion Prevention System.

JSON Javascript Object Notation.

KLD Kullback-Leibler Divergence.

LDA Latent Dirichlet Allocation.

LDAP Ligthweight Directory Access Protocol.

ML Machine Learning.

RAM Random Access Memory.

SIEM Security Information and Event Management.

SNMP Simple Network Management Protocol.

SOC Security Operations Center.

SPAN Switched Port Analyzer.

TAP Terminal Access Point.

TIC Tecnologias de Informação e Comunicação.

UBA User Behavior Analytics.

UEBA User and Entity Behavior Analytics.

URL Uniform Resource Locator.

VPN Virtual Private Network.

WEC Windows Event Collector.

WEF Windows Event Forwarding.

Capítulo 1

Introdução

1.1 Motivação

Nos dias que correm, as organizações revelam uma preocupação crescente com todas as questões que envolvem o domínio da cibersegurança nos seus ambientes. A Altice Portugal, empresa de referência na área das Tecnologias de Informação e Comunicação (TIC), procura endereçar estas preocupações e, para tal, possui um *Cyber Security Operations Center* (CSOC) e uma *Computer Security Incident Response Team* (CSIRT), tanto para uso interno, como uso externo. O principal objetivo do CSOC é monitorizar eventos de segurança relacionados com os ativos da organização, identificando incidentes de segurança. Uma ferramenta fundamental para cumprir com este desiderato são os sistemas de segurança da informação e gestão de eventos (*Security Information and Event Management* - SIEM) [8, 9]. Estes permitem recolher eventos originados em diversas fontes, normalizá-los num formato comum, guardá-los para análise forense, e correlacioná-los para identificar atividades maliciosas em tempo-real.

Por outro lado, com a cada vez maior sofisticação e capacidade de camuflagem dos ataques realizados às organizações, evitando os padrões/assinaturas conhecidos, tem sido explorado um novo tipo de vulnerabilidade: a ameaça interna. Na realidade, este tipo de ameaça já não é novo - o que é novo é o seu crescimento no domínio da cibersegurança. Segundo Allen *et al.* [10], a existência de um trabalhador malicioso pressupõe uma deterioração progressiva das componentes cognitivas e comportamentais de um trabalhador confiável, isto é, existe uma fita de tempo para a prossecução de atos maliciosos. Como tal, a criação e estabelecimento de perfis de comportamento base para cada utilizador pode revelar-se uma mais valia, possibilitando detetar anomalias comportamentais que se traduzam em ameaças e/ou incidentes de segurança. É neste contexto que surge o conceito de *User Behavior Analytics* (UBA), ou ainda mais recentemente, a agregação deste com outras quaisquer entidades do mundo digital, *User and Entity Behavior Analytics* (UEBA).

Se no início as soluções UEBA eram disponibilizadas de forma isolada, nos últimos anos tem-se assistido a uma integração destas nos SIEM, no sentido de estes tirarem partido das capacidades analíticas avançadas das ferramentas UEBA [3]. A ferramenta SIEM utilizada pela Altice Portugal enquadra-se nesta situação. Contudo, ainda existe um longo caminho a percorrer no que toca à validação da utilização de abordagens comportamentais para identificar e mitigar a dita ameaça interna [10]. Este projeto endereça esta limitação, procurando provar a capacidade da plataforma SIEM utilizada na Altice Portugal na deteção e análise de anomalias comportamentais de segurança em utilizadores desta organização.

1.2 Objetivos

O principal objetivo deste projeto passa por provar a exequibilidade de implementação de uma solução, de forma eficaz e eficiente, para deteção e análise de anomalias comportamentais de segurança dos utilizadores de uma grande organização, com impacto na cibersegurança dessa mesma organização, em tempo quase real. Para tal, faz-se uso da plataforma *IBM QRadar* utilizada pelo CSOC da Altice Portugal, nomeadamente da aplicação UBA, recorrendo às funcionalidades de *Machine Learning* por esta disponibilizadas. Esta capacidade de deteção e análise é fundamental, pois permite prevenir eventuais ataques ou ações negligentes de utilizadores confiáveis da organização. Embora a plataforma utilizada tenha capacidade de suportar a monitorização interna da Altice Portugal, e a monitorização externa de outras entidades, nomeadamente clientes empresariais que contratam esses serviços, o âmbito deste projeto restringe-se apenas aos utilizadores da Altice Portugal.

Neste contexto, torna-se necessário compreender a aplicação UBA, realizando uma análise técnica crítica à sua aplicabilidade efetiva na deteção e análise de casos de uso concretos relevantes para a gestão da cibersegurança da Altice Portugal, nas áreas dependentes do comportamento dos seus utilizadores. Para atingir este objetivo, importa definir e priorizar, primariamente, uma lista de casos de uso que endereçam as principais preocupações ao nível das possíveis ameaças internas, nomeadamente no que se refere ao possível abuso de credenciais e à existência de discrepâncias tempo-espaciais. Posteriormente, interessa validar quais as fontes que contribuem com os dados necessários para a monitorização dos casos de uso anteriormente identificados. Adicionalmente, considera-se também importante comparar a capacidade de deteção dos modelos analíticos disponibilizados por omissão, pela aplicação, com os modelos analíticos personalizados para os casos de uso identificados.

Por fim, importa garantir também que a solução implementada apresenta um desempenho que não compromete o atual normal funcionamento da referida plataforma *IBM QRadar*.

1.3 Contribuições

A principal contribuição deste projeto passa pela verificação e validação da exequibilidade de implementação de uma solução que contribua para uma deteção e análise, mais rápida e antecipada, de anomalias comportamentais de segurança dos utilizadores que estejam a ocorrer na rede, visando assim proporcionar uma maior visibilidade, capacidade de resposta e resiliência à ocorrência de incidentes. Deste modo, é possível identificar as seguintes contribuições:

- Criação e definição de uma metodologia robusta e escalável que permita caraterizar as necessidades de monitorização específicas de uma grande organização, nomeadamente na identificação, especificação e priorização de uma lista de casos de uso, bem como na escolha e identificação das fontes de dados que contribuem para uma caraterização mais precisa dos casos de uso anteriores;
- Análise comparativa crítica quanto à efetiva capacidade de deteção de anomalias comportamentais entre os modelos analíticos disponibilizados por omissão pela aplicação ML e os modelos personalizados criados na mesma;
- Estudo e avaliação dos parâmetros, e respetivos valores, dos modelos analíticos que permitem maximizar a taxa de deteção de anomalias comportamentais, enquanto se minimiza a taxa de falsos positivos.

Em suma, este projeto visa permitir compreender a real capacidade de deteção de anomalias comportamentais de segurança dos utilizadores, fazendo uso da plataforma *IBM QRadar*, mais concretamente das suas aplicações UBA e ML, contribuindo assim com informação valiosa para melhorar o panorama

global de monitorização de segurança.

1.4 Estrutura do documento

Este documento está organizado da seguinte forma:

- Capítulo 1 - É o presente capítulo deste documento, o qual apresenta em traços gerais a motivação, os objetivos que são propostos alcançar, as principais contribuições deste trabalho e a estrutura do documento deste projeto;
- Capítulo 2 - Descreve um conjunto de conceitos importantes no domínio da deteção e análise comportamental de utilizadores, revendo posteriormente que trabalhos têm sido desenvolvidos nesta área de conhecimento e estudo. Neste capítulo inclui-se também uma descrição da ferramenta que será utilizada ao longo deste projeto;
- Capítulo 3 - Aborda a análise do problema que levou à elaboração deste projeto, bem como o desenho da solução, nomeadamente na definição das fontes de dados e casos de uso a implementar;
- Capítulo 4 - Descreve o processo de implementação da solução proposta, bem como os diferentes tipos de testes que permitiram realizar uma adequada avaliação da mesma;
- Capítulo 5 - Este capítulo efetua a análise e discussão do trabalho desenvolvido, com base nos resultados obtidos a partir dos diferentes testes realizados;
- Capítulo 6 - Neste último capítulo são sumarizadas as principais conclusões deste projeto, bem como é analisado o cumprimentos dos objetivos propostos no início deste trabalho. Simultaneamente, são apresentadas também as oportunidades de trabalho futuro que decorrem no seguimento deste projeto.

Capítulo 2

Trabalho Relacionado

2.1 Contexto

2.1.1 Ameaça Interna

Ameaça interna pode ser definida como qualquer atividade ou ação, intencional ou negligente, realizada por utilizadores "internos confiáveis"da organização, da qual resultam perdas na confidencialidade, integridade e disponibilidade na informação e/ou infraestruturas tecnológicas da organização [11, 12]. Ações como roubo de informação, sabotagem, fraude e/ou espionagem recaem na categoria de ameaça interna, sendo que regra geral, qualquer destas requer abuso de privilégios de acesso.

A ameaça interna tem emergido na segurança das organizações e tem vindo a receber uma maior atenção nos últimos anos. Um estudo realizado pela *Cibersecurity Insiders* [13] revela que 53% das organizações foram alvo de ataques internos no último ano, sendo que 90% das mesmas se sentem vulneráveis a ataques provenientes desta origem. Já em 2016, segundo a *IBM* [14], 60% dos incidentes de segurança tiveram origem em utilizadores internos confiáveis. Para além destes dados, o estudo revela também que tanto os trabalhadores com privilégios normais (56%), como os trabalhadores com privilégios de administração (55%), seguidos pelos trabalhadores temporários/contratados/prestadores de serviços (42%) se traduzem nos maiores riscos de segurança internos para as organizações. Ativos como as bases de dados e os servidores de ficheiros continuam no topo dos ativos que apresentam um risco maior [13, 11].

Deste modo, podem existir três tipos de ameaças internas, para os quais as organizações necessitam de olhar e compreender [4]:

- Trabalhadores negligentes/não intencionais - este tipo de ameaça deriva de trabalhadores sem intenção maliciosa, os quais não seguem os procedimentos e normas da política de segurança da organização, ou causam erros inadvertidamente;
- Trabalhadores maliciosos/mal intencionados - corresponde a um trabalhador que tem ou tinha acesso à rede, sistemas e dados da organização e, de forma deliberada e intencional, excedeu ou fez mau uso do seu acesso, prejudicando gravemente a confidencialidade, integridade e disponibilidade da informação ou dos sistemas de informação [11]. Motivações como descontentamento ou oportunismo financeiro estão na principal origem desta ameaça;
- Trabalhadores comprometidos - reflete o caso de um trabalhador cujas credenciais foram roubadas ou cujo dispositivo foi infetado com *malware*, geralmente com recurso a ataques de *phishing*. A partir daqui os atacantes podem pesquisar por partilhas de ficheiros, escalar privilégios e infetar outras máquinas até aceder aos ativos mais importantes, conforme apresentado na figura 2.1.

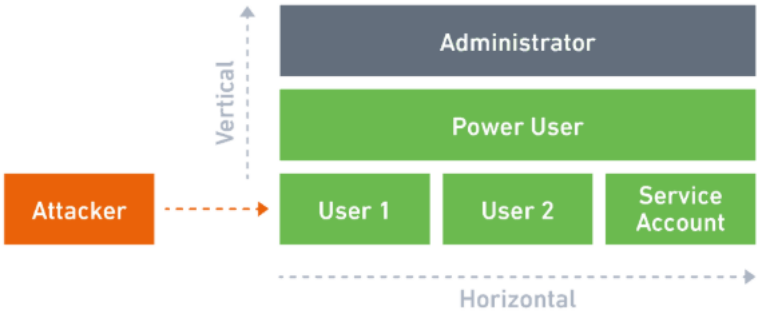

Figura 2.1: Escalonamento de privilégios (extraído de Matthews [1])

Apesar de já existir o esforço, de há uns anos a esta parte, para detetar e prevenir este tipo de ameaças, continuam-se a registar perdas significativas. Tal deve-se sobretudo aos seguintes fatores: 1) as soluções existentes não prestam atenção suficiente a indicações prematuras, e só emitem um alerta quando já foi causado algum tipo de danos; 2) a grande maioria das soluções existentes assenta numa fonte individual de auditoria; 3) os métodos analíticos dependem muito do domínio de conhecimento para extrair caraterísticas ou estabelecer regras, o que se traduz numa capacidade limitada para lidar com novas ameaças [12]. Por outro lado, este tipo de ataques está a revelar-se cada vez mais difícil de encontrar e lidar porque: 1) a existência de um evento anómalo não significa que seja obrigatoriamente malicioso, e pequenas alterações ocorrem todos os dias, o que pode originar um número elevado de falsos positivos, dispersando a atenção e levando à descrença pelos analistas; 2) um evento, por si só, pode não ser suficiente para comprovar que um ataque está a ocorrer; 3) os sinais dados pelos atacantes são cada vez mais subtis, atuando em passos curtos, de forma a evitar o comportamento de ataques conhecidos [4]. Para finalizar, Liu *et al.* [12] afirma que as principais consequências que resultam de um ataque originado internamente são: exfiltração de dados, violações da confidencialidade, integridade e disponibilidade dos dados e sabotagem dos sistemas de informação.

2.1.2 *Security Analytics*

O ritmo dos ataques no contexto da cibersegurança está a crescer de dia para dia. As ameaças a que as organizações estão sujeitas podem ser definidas segundo uma grelha de duas dimensões, conforme ilustra a figura 2.2: ataques, sejam eles conhecidos ou não, e atacantes, sejam também eles conhecidos ou não. Os ataques conhecidos podem ser detetados usando regras estáticas, verificando assinaturas ou padrões através de sistemas antivírus, sistemas de prevenção de intrusões (*Intrusion Prevention Systems* - IPS), *firewalls*, aplicacionais ou não, e SIEM. Quando são conhecidos os atacantes, os seus atributos, como é o caso dos endereços IP, URL's e ficheiros, podem ser bloqueados e/ou colocados sob observação. Por outro lado, quando os atacantes são conhecidos através de fontes de *intelligence* mas os seus métodos são desconhecidos, pode ser utilizada uma plataforma de *Big Data* para recolher os fluxos de tráfego de rede, *proxies* e acessos de utilizadores, para correlacionar com quaisquer Indicadores de Compromisso (*Indicators of Compromise* - IoC) relacionados com os atacantes. Contudo, todas estas abordagens anteriores de regras, assinaturas e fontes de *intelligence* falham quando se pretende detetar ataques desconhecidos[1], originados por atacantes não conhecidos. A necessidade emergente e crítica por *Security Analytics* surge então da necessidade de deteção de ataques desconhecidos. Tal não significa que não possa trazer algum valor acrescido aplicá-la aos restantes quadrantes, no entanto, deverá ser sempre considerado segundo o rácio custo-benefício onde, se uma estratégia baseada em regras é insuficiente, então o próximo passo será definir um caso de uso apropriado e aplicar *Security Analytics* [2].

[1] Vulgarmente designados por *Zero Day Attacks*.

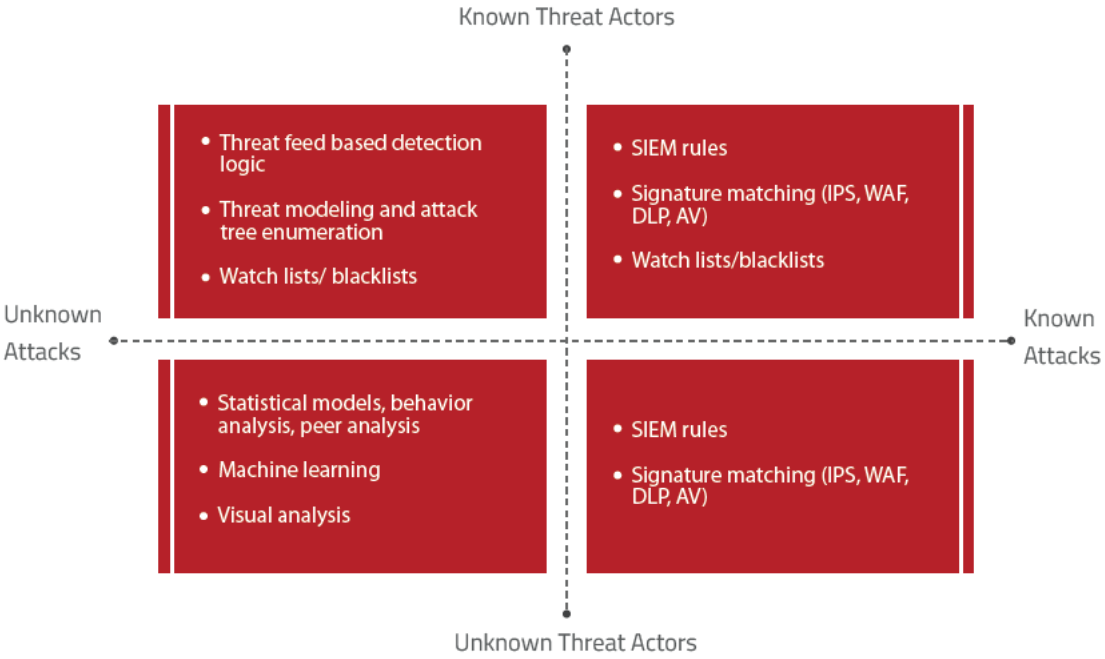

Figura 2.2: Grelha de ataques vs atores (extraído de Vasudevan [2])

É pois neste domínio de maior complexidade, sofisticação e camuflagem dos ataques que surge este conceito de *Security Analytics*. Neste contexto, Shackleford [15] afirma que "*Nós precisamos de mais dados de mais fontes diferentes, sobre períodos de tempo mais longos, para realmente desenvolver uma melhor compreensão do que está a acontecer no ambiente de rede...e deve ser pesquisada uma forma de comportamento anormal, mas eles (equipas de segurança) não sabem o que procuram, nem como captar*". Mormente, segundo Shackleford [15], *Security Analytics* pode ser definido como a análise a produzir sobre um grande conjunto de dados, de fontes distintas, recorrendo a tecnologias e técnicas que permitem correlacionar e reportar eventos e/ou padrões de interesse, que podem indicar comportamento malicioso no ambiente, de uma forma precisa e rápida. Da mesma forma, é possível dizer que o principal objetivo da *Security Analytics* é disponibilizar capacidades preditivas para além dos métodos tradicionais para prever ameaças futuras, sendo capazes de desenvolver perfis de comportamento base, baseando-se no comportamento atual e passado, que permitam detetar padrões novos e não usuais, tanto para ataques conhecidos como desconhecidos. Chuvakin e Barros [16] definem *Security Analytics* como uma análise avançada de um conjunto de dados para atingir um resultado útil no domínio da segurança, isto é, utilizando outros métodos que não verificação de regras ou estatísticas básicas, por exemplo *Machine Learning* mas não só, analisar qualquer tipo de dados desde *logs*, fluxos de tráfego de rede, sessões, transações e outros, para detetar potenciais atividades maliciosas. Já segundo Pritz [17], o conceito de *Security Analytics* assenta na descoberta, interpretação e comunicação de padrões significativos em grandes conjuntos de dados, fazendo uso de aprendizagem automática e inteligência artificial, e com capacidade de automatizar e tornar o processo mais inteligente e eficiente. Mais, afirma ainda que existem quatro definições chave que contribuem para melhor definir este conceito. Estas definições são: perfil de comportamento - onde se procura medir o comportamento normal para reconhecer quando o comportamento foge desse padrão; análise de grupo de pares - através da agregação de pares com atributos em comum e comparação das ações de um elemento com o padrão do grupo; contexto de negócio e ameaça importantes - para definir a valorização do risco a atribuir; e modelação da ameaça - para permitir prever e priorizar a investigação e a resposta. Por outro lado, Vasudevan [2] refere que o ponto de início da *Security Analytics* deveria ser identificar os riscos que não podem ser monitorizados com os métodos

tradicionais, definindo depois casos de uso para os acompanhar.

2.1.3 Deteção de Anomalias e Comportamento

"*Deteção de anomalias é descrito como o problema de encontrar padrões nos dados cujo comportamento não está de acordo com o esperado. Os padrões não conformes são geralmente designados como anomalias ou como pontos isolados em diferentes domínios de aplicação*" [18]. Conforme indicado por Ma [19], em deteção de anomalias existem quatro desafios principais:

- definir um conjunto que inclua todas as diferentes possibilidades de comportamento normal não é viável;
- produzir uma fronteira clara e precisa entre comportamento normal e anomalia é uma tarefa difícil;
- possuir um conjunto de dados devidamente categorizados para treino e classificação de anomalias é dispendioso;
- aplicar técnicas num domínio, desenvolvidas num outro contexto não é linear, uma vez que o comportamento malicioso é dinâmico.

Nesta sequência, importa também definir o que se entende por comportamento do utilizador. De acordo com a biologia, comportamento são as respostas internas coordenadas, de todos os organismos vivos, a todos os estímulos internos ou externos. De igual modo, comportamento digital será tudo aquilo que é realizado no mundo digital [20]. Assim, para melhor definir o comportamento de um utilizador, devemos identificar a granularidade temporal da análise (hora, dia, semana, etc.), bem como o conjunto de caraterísticas que permitem caraterizar o padrão de acesso, em cada período temporal, para cada utilizador [21].

2.1.4 *User Behavior Analytics (UBA) e User and Entity Behavior Analytics (UEBA)*

A melhor forma de habilitar as organizações a detetar estas ameaças internas silenciosas é analisando o comportamento dos utilizadores, e procurando por anomalias que permitam fazer uma deteção precoce, de forma a conter e evitar os ataques [14]. UBA pode ser definido como o processo de cibersegurança capaz de detetar ameaças internas, ataques direcionados ou fraudes financeiras, fazendo uso de métodos analíticos e algoritmos para construir padrões de comportamento *standard* de utilizadores ou grupos de utilizadores, ao longo do tempo, possibilitando assim detetar atividade que é anómala a estes mesmos padrões. Face aos desenvolvimentos na tecnologia UBA, tornou-se necessário evoluir este conceito para UEBA. A expansão deste conceito passou então a incluir o termo de entidade, onde se incluem todos os dispositivos, aplicações, servidores, repositórios de dados, isto é, qualquer entidade que possua um endereço IP, pois é considerado que estas entidades podem ter um papel fundamental na descoberta de atividade maliciosa não visível e imperceptível para os sistemas de monitorização, como SIEM ou um sistema de prevenção de perda de dados (*Data Loss Prevention* - DLP) [3]. UBA recolhe então vários tipos de dados como a estrutura organizacional, papéis de utilizadores, responsabilidades do cargo, registo de atividade e localização geográfica, uma vez que os seus algoritmos de análise consideram fatores como a informação de contexto, atividades contínuas, duração das sessões e atividades de grupos de pares idênticos para comparar comportamentos. UBA determina então uma linha base daquilo que é o comportamento normal de uma dada entidade individual, ou de um dado grupo de entidades, de acordo com o histórico de dados e atividade. Qualquer desvio das atuais atividades do utilizador, quando comparadas com o comportamento passado normal, será significativo se o utilizador se comportar anormalmente [19, 11]. Quaisquer eventos anormais são pois detetados e agregados segundo uma valorização, a qual permite fornecer um valor de risco integrado para cada entidade, podendo até ser

utilizadas outras ferramentas de segurança que contribuam para este processo. As entidades que representam um risco mais elevado serão sinalizadas a um analista que deverá investigar aquele particular comportamento, no contexto do papel e responsabilidades dessa entidade na organização.

Nestes últimos anos têm sido desenvolvidos diversos contributos, procurando criar modelos de comportamento anómalo. Neste sentido, descreve-se de seguida o principal trabalho desenvolvido neste domínio:

- Salem e Stolfo [22] desenvolveram um método que agrupa comandos *Unix* e as aplicações do *Windows* em categorias, a partir dos *logs* de auditoria, e aplicando aprendizagem automática para detetar ações maliciosas;
- Angeletou *et al.* [23] fazem uso de análise estatística combinada com um modelo semântico e regras para representar e computar o comportamento observado nas comunidades *online*;
- Hu *et al.* [24] propõem um método que assenta na utilização dos meta-eventos (normalização) das atividades dos utilizadores, combinando todas as fontes disponíveis para detetar se os seus padrões de comportamento, atualmente observados, são consistentes ou não com os registados no passado. Cada utilizador é processado de forma independente dos restantes;
- Thompson [25] propõe que sejam desenvolvidos um conjunto de modelos, seguindo uma abordagem baseada no conteúdo, para detetar eventuais sinais de má utilização. Posteriormente, os modelos são comparados entre si para detetar eventuais discrepâncias, ou são pesquisadas assinaturas conhecidas de má utilização, de forma individual;
- Maloof e Stephens [26] constroem detetores, um por cada tipo de atividade (por exemplo: pesquisa, download, impressão, navegação na Internet, etc) usando depois redes *Bayesianas* para definir a ordenação. Eldardiry *et al.* [27] também constroem detetores, mas um por cada tipo de domínio de dados (por exemplo: autenticação, email, web, etc.) e procura anomalias antes de fundir os resultados;
- Tanto a *IBM* como a *Splunk* [28, 29] constituem-se como soluções UBA, disponibilizadas por empresas de referência que constroem indicadores baseados em perfis estatísticos para encontrar anomalias em eventos obtidos a partir de *logs* ou outro tipo de fontes;
- Xiangyu *et al.* [11] propõem uma plataforma de análise de comportamento de utilizadores com capacidade para recolher e pré-processar os *logs* dos sistemas e aplicações, extrair e agregar as caraterísticas, gerando um vetor para cada utilizador, e, fazendo uso de um conjunto de vários algoritmos não supervisionados, detetar padrões fora do normal;
- Shashanka *et al.* [21] descrevem uma solução UEBA implementada pela plataforma *Niara Security Analytics*, cuja principal ideia assenta no conceito de uma entidade 360, que reúne toda a informação (por exemplo: utilizador, endereço IP, dispositivo, etc.) obtida a partir de diversas fontes, a qual depois de combinada com métodos e algoritmos de aprendizagem automática, permite gerar um perfil de risco para essa entidade;
- Lashkari *et al.* [30] propõem um modelo de perfil de utilizador, o qual cobre todas as fontes disponíveis e caraterísticas relacionadas, ultrapassando assim as deficiências de modelos anteriores.

Para além deste trabalho desenvolvido, e apesar de o mesmo revelar que existem muitos traços comuns entre os diferentes trabalhos, segundo Ware *et al.* [31], é de consenso geral que hoje em dia, uma solução UEBA para ser bem sucedida necessita de apresentar três caraterísticas fundamentais:

- Preditiva - cada vez mais as organizações precisam de sistemas que tenham a capacidade de prever e alertar de forma antecipada um conjunto grande de eventos adversos, recorrendo a

todos os dados internos e externos disponíveis, ao mesmo tempo que a taxa de falsos positivos diminui consideravelmente. Para tal, torna-se necessário construir um modelo, de acordo com as suas especificidades de segurança, que corra em tempo real, com todas as fontes de dados relevantes, e que seja capaz de gerar alertas quando são detetados valores de risco elevados;

- Adaptável - o sistema deve ser capaz de se ajustar às especificidades de cada organização, bem como de evoluir e compreender as ameaças ao longo do tempo. Mais, deve ser flexível e aberto para cumprir com a política de segurança e *standards* da tecnologia;
- Escalável - o sistema deve ser capaz de absorver novos conjuntos de dados, independentemente do volume ou formato, assim que estes estejam disponíveis, e sempre sem comprometer o atual desempenho, nem produzir falsos positivos adicionais.

Conforme referido por Sadowski *et al.* [3], os sistemas UEBA assentam em três pilares fundamentais ou podem ser caraterizados segundo três dimensões: casos de uso, métodos analíticos e fontes de dados. De seguida, cada um destes pilares é descrito de forma mais detalhada.

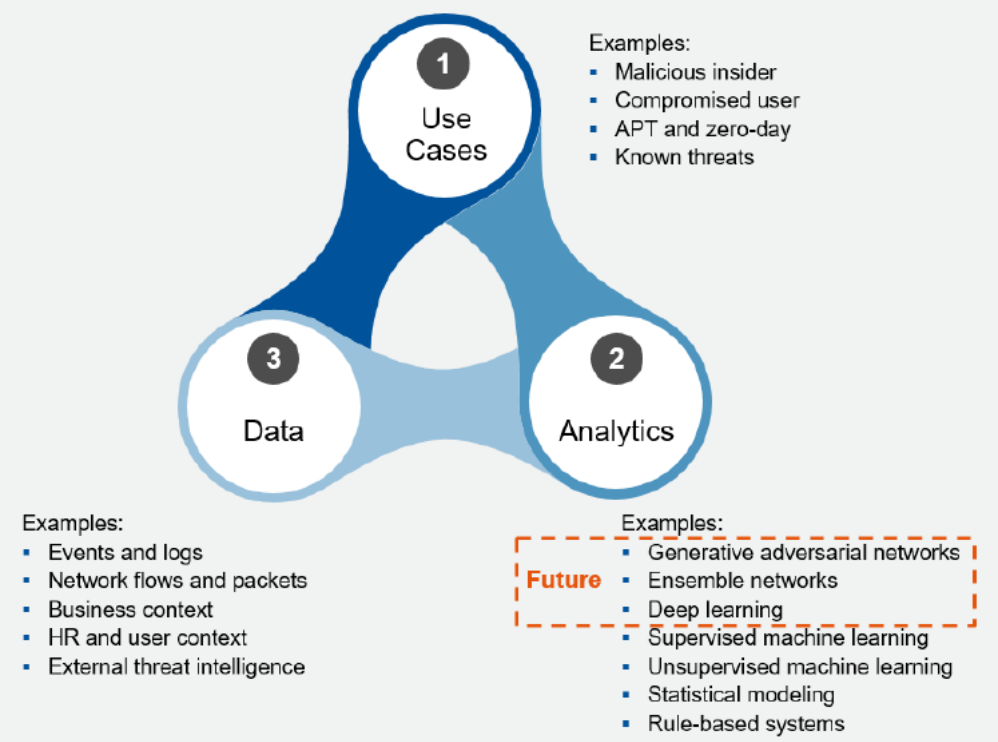

Figura 2.3: Pilares UEBA (extraído de Sadowski et al. *[3])*

2.2 Casos de Uso

Os casos de uso são originalmente associados ao processo de desenvolvimento de *software*, mas têm vindo a ser adotados no domínio da cibersegurança devido ao crescimento e evolução dos SIEM. Estes procuram refletir o tipo de ataques que se pretende prevenir, ou as etapas da cadeia de processos de negócio que se pretende monitorizar. De acordo com Chuvakin e Barros [32], casos de uso podem ser definidos como uma "*condição ou evento específico, geralmente relacionados com uma ameaça concreta, o qual deverá ser detetado e reportado por uma ferramenta de segurança*". No contexto de UEBA, o objetivo dos casos de uso passa por fornecer dados sobre o comportamento dos utilizadores e entidades, bem como efetuar a monitorização, deteção e alerta de eventuais comportamentos anómalos,

não sendo aconselhável que o foco resida apenas num único caso de uso [3]. Mormente, tem-se assistido que cada vez mais as soluções UEBA suportam uma variedade de casos de uso. Sadowski *et al.* [3] afirma que casos de uso como a monitorização de acessos e movimentos de dados não autorizados, atividades suspeitas de utilizadores privilegiados, atividades maliciosas ou não autorizadas dos empregados, acesso e utilização não usual de recursos na *cloud* continuam a ser dos mais utilizados. Por outro lado, existe também, tipicamente, um conjunto de casos de uso, nomeadamente deteção de fraude e monitorização de empregados, não centrados em cibersegurança, para os quais estas ferramentas UEBA também são utilizadas. Contudo, este tipo de casos de uso requer, normalmente, fontes de dados não relacionados com segurança ou necessitam de modelos analíticos específicos.

De forma mais detalhada, Sadowski *et al.* [3] referem que os cinco principais domínios de casos de uso em que, tanto as soluções UEBA, como os seus utilizadores estão de acordo são:

- Utilizador Interno Malicioso - permite monitorizar os trabalhadores da organização ou os parceiros externos confiáveis, procurando comportamentos não habituais, maliciosos ou abusivos, que visam infligir danos na organização, isto é, está orientado para encontrar utilizadores internos envolvidos em atividades maliciosas, e não ameaças avançadas onde contas internas são comprometidas. Geralmente, não são monitorizadas contas de serviço ou de entidades não humanas. Considerando que uma intenção maliciosa é difícil de identificar, torna-se necessário analisar a informação de contexto não estruturada, como conteúdo de correio eletrónico, indicadores de desempenho e informação das redes sociais, para determinar e caraterizar o comportamento contextual do trabalhador;
- Utilizador Interno Comprometido - destina-se a detetar e analisar atividades maliciosas de atacantes que se infiltram nas organizações, e se movem lateralmente por toda a infraestrutura de tecnologias de informação. Ameaças Persistentes Avançadas (*Advanced Persistent Threat* - APT) e ameaças desconhecidas, designadas muitas vezes por
  iiataques dia-zero, escondidas habitualmente atrás de utilizadores legítimos ou contas de serviço, situam-se entre as principais ameaças e são de deteção bastante difícil. No entanto, estas ameaças avançadas, apesar de não terem uma assinatura conhecida, acabam também por forçar os ativos da organização a ter um comportamento distinto daquele que apresentam habitualmente, geralmente associado a utilizadores e entidades não suspeitas;
- Roubo de Dados - destina-se a prevenir a perda de dados nas organizações, e confere a possibilidade de detetar, tanto exfiltração originada por utilizadores internos, como por atacantes externos. Neste tipo de casos de uso, torna-se importante a adição de contexto, como tráfego de rede (por exemplo, *web proxy*) e dados dos dispositivos terminais, para melhor deteção;
- Gestão de Identidades e Privilégios de Acesso - permite monitorizar e analisar o comportamento dos utilizadores, com o objetivo de identificar acessos fora do normal, existência de privilégios excessivos e até limpar contas não utilizadas/inativas. Este tipo de caso de uso aplica-se a todo o tipo de utilizadores e contas, incluíndo utilizadores com privilégios especiais (por exemplo, administradores) e contas de serviço;
- Priorização de Incidentes - possibilita ajudar a organização na priorização de incidentes ou potenciais incidentes, ou seja, recorrendo a modelos de ameaça e padrões de utilização conjugados com conhecimento sobre a estrutura da organização, as soluções UEBA conseguem determinar que incidentes serão mais perigosos ou significativamente fora do normal, atribuindo um fator de ponderação superior a estes.

2.3 Métodos Analíticos

Os métodos analíticos, como é o caso das regras, são intuitivos e de fácil concepção e implementação e, geralmente, são independentes e definidas por cada fonte de dados. Mais, a sua correlação não procura dados para além de um pequeno período de tempo, e não são capazes de abordar coletivamente a heterogeneidade dos dados [24]. As regras são determinísticas por natureza logo, quando uma regra é despoletada, será sempre gerado um alerta. No entanto, esta abordagem tradicional e antiga, tal como a baseada em padrões, já não é suficiente, pois só permite defender daquelas ameaças ou ataques conhecidos, isto é, falhará em descobrir ataques mais sofisticados. Neste sentido, emergem os métodos analíticos mais avançados, como é o caso da aprendizagem automática, vulgo *Machine Learning*, os quais procuram afastar-se do estado binário (ameaça / não ameaça) e apresentar um valor de risco para cada utilizador ou ativo, baseado nos modelos e contexto da organização. Estes métodos avançados, por sua vez, são heurísticos onde os modelos computam constantemente a probabilidade de um evento ser anómalo, o nível de anormalidade do evento e a probabilidade de o mesmo ser uma ameaça. Por outro lado, importa também referir que estes métodos são complementares dos métodos tradicionais, e não mutuamente exclusivos, uma vez que existe bastante valor acrescentado em fazer uso de assinaturas e regras para detetar ataques conhecidos e/ou definir níveis de atividade não desejada numa relação custo-benefício vantajosa [3]. Os diferentes métodos analíticos possibilitam então a deteção de anomalias, devendo os mesmos ser escolhidos de acordo com os casos de uso implementados. Estes métodos têm capacidade de aprender o comportamento normal das diferentes entidades e grupos de entidades, identificando posteriormente qualquer desvio da normal atividade. Quando existir esse desvio, o sistema adiciona o valor de risco dessa atividade à respetiva entidade. Quanto maior for o nível de risco da atividade, ou a sua fuga ao padrão de comportamento normal, maior será o seu valor de risco. À medida que mais e mais comportamentos suspeitos vão sendo acrescentados, o valor de risco vai aumentando até atingir um limite bem definido, altura em que um analista é notificado [33]. Existe, assim, uma agregação de vários eventos, os quais contribuem para o valor de risco de um utilizador, ganhando o analista uma visão de alto nível, e facilitando a deteção de ataques mais elaborados.

Aprendizagem automática pode ser definida como um tipo de inteligência artificial que fornece aos computadores a capacidade de aprender, sem serem explicitamente programados para cada cenário diferente. Apesar de a aprendizagem automática não ser um novo conceito ou tecnologia, a sua aplicação no domínio da cibersegurança é. Segundo Shackleford [34], num estudo realizado em 2016, apenas 22% das organizações utilizam aprendizagem automática. Contudo, face aos constantes novos desafios, perspetiva-se um crescimento mais rápido, bem como que venha a ser definido como uma boa prática neste contexto.

As soluções atuais de aprendizagem automática estão baseadas num conjunto bem documentado de modelos matemáticos, os quais recebem um conjunto de parâmetros e caraterísticas (elementos individuais relevantes dos dados) originários dos algoritmos, e aplicam cálculos específicos. Enquanto os algoritmos básicos e processos não são secretos, a forma como os mesmos são usados e estão implementados determina o real valor que cada uma das soluções acrescenta. Por outro lado, e sem qualquer surpresa, os modelos podem tornar-se complexos, e bastante consumidores de recursos e, como um único modelo não é passível de ser aplicado a todos os casos de uso, a escolha do mesmo e do conjunto de dados associado revela-se crítica e fundamental. Assim, para a deteção de anomalias, podem ser utilizadas as seguintes variantes de algoritmos de aprendizagem automática:

- Supervisionados - esta abordagem requer que os modelos sejam ensinados e treinados para detetar uma determinada condição específica. O sistema é alimentado com um conjunto de treino, onde estão devidamente identificados os bons e os maus comportamentos. Depois de treinados, esses comportamentos podem então ser detetados, mesmo em conjuntos de dados nunca antes

observados. As desvantagens desta variante decorrem da necessidade de categorizar o conjunto de dados, o que faz com que seja suscetível a erros, consumidor de tempo e dispendioso;

- Não Supervisionados - nesta abordagem não é requerido qualquer treino ou programação *à priori*. O modelo aprende a partir de conjuntos de dados não categorizados, procurando identificar grupos com atributos similares, *clustering*, e construir as linhas base do que é expectável ser normal, o que dependerá das distâncias, densidade e estatísticas dos dados. Este tipo de abordagem assume de forma implícita que as instâncias normais representam a maioria dos dados não categorizados. Caso tal não suceda, e existam mais instâncias anormais do que normais, o modelo sofrerá com uma elevada taxa de falsos positivos. Desta forma, pode afirmar-se que o modelo categorizará as instâncias como normais ou anormais, ou ainda, por outras palavras, atribuirá um valor de anomalia, pois ele não sabe o que é um bom ou um mau comportamento, cabendo sempre a interpretação a um analista especializado;
- Semi-supervisionados - consiste numa abordagem híbrida que faz uso de modelos não supervisionados, e onde os alertas gerados realimentam o próprio modelo, de modo a baixar a taxa de falsos positivos e aumentar a confiança nos resultados. Como ponto negativo, este treino do sistema em tempo real faz com que os modelos demorem mais tempo a ficar eficientes na deteção de comportamentos desviantes;
- *Deep Learning* - possibilita a triagem e investigação de alertas virtuais. O sistema treina em conjuntos de dados que representam alertas de segurança e nos seus resultados da triagem, realiza a auto-identificação de caraterísticas, e é capaz de prever resultados de triagem para novos conjuntos de alertas de segurança. Segundo Sadowski *et al.* [3] este tipo de métodos analíticos constituirá o futuro da utilização de *Machine Learning*;
- Modelos *Ensemble* - abordagem na qual diferentes metodologias e modelos podem correr concorrentemente uns com os outros, e onde cada um produz um voto, de modo a procurar obter um melhor desempenho preditivo. Contudo, esta abordagem requer capacidades de computação mais exigentes do que as requeridas pela utilização de um único algoritmo em tempo real. Tal como a abordagem anterior, segundo Sadowski *et al.* [3], este tipo de métodos analíticos constituirá o futuro da utilização de *Machine Learning*.

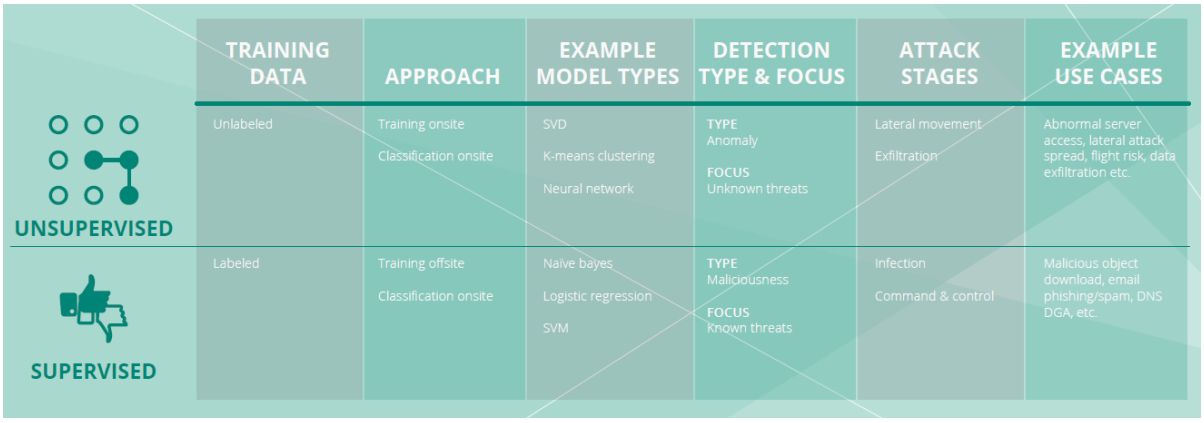

Figura 2.4: Comparação entre Modelos Supervisionados e Não Supervisionados (extraído de Aruba Networks - CISO Guide[4])

Como afirma Sadowski *et al.* [3] podem também ser definidas duas estratégias quanto à forma como os utilizadores podem interagir com os diferentes métodos analíticos que as soluções UEBA existentes no mercado disponibilizam:

- Fechadas - os métodos analíticos não podem ser vistos nem modificados pelo utilizador, e os compradores devem aceitar os casos de uso e modelos disponibilizados;
- Abertos - estes produtos expõem os elementos aos utilizadores. As ferramentas com métodos analíticos mais básicos tendem a ser abertos, pois baseiam-se sobretudo em regras, assinaturas e padrões que requerem configuração para os tornar utilizáveis. As que fazem uso de analíticas mais avançadas, com aprendizagem automática associada, podem permitir aos utilizadores alterar algumas das variáveis ou atributos nos seus modelos;

Para finalizar, importa tecer algumas considerações relativamente a alguns cuidados a ter com a criação de perfis base e, consequentemente, deteção de anomalias. Em primeiro lugar, é preciso salientar que o comportamento de utilizadores com privilégios especiais, ou outros, pode ser altamente irregular dependendo das suas funções, tornando assim difícil a definição do comportamento base, e originando problemas na deteção de anomalias. Em segundo lugar, há que considerar também que um dado utilizador ou grupo pode apresentar um comportamento desviante logo desde o início do estabelecimento do perfil base, o que fará com que, mais tarde, esse mesmo comportamento desviante não seja detetado, pelo que poderá fazer sentido definir a periodicidade de estabelecimento de novos perfis. Por último, a deteção de comportamento suspeito de utilizadores com privilégios especiais, programadores ou utilizadores comprometidos, não pode assentar apenas nos métodos fornecidos pelos fabricantes de soluções UEBA, uma vez que os mesmos não estão ainda suficientemente provados e testados. Como tal, as organizações deverão ficar responsáveis pela definição, inclusão e afinação de casos de uso apropriados ao seu contexto para funcionarem, lado a lado, com os métodos disponibilizados [3].

2.4 Fontes de Dados

O terceiro pilar dos sistemas UEBA são as fontes de dados. Quanto melhor for a qualidade dos dados fornecidos por estas, melhores serão os resultados que serão obtidos, ou seja, quanto maior for a confiabilidade dos dados, melhor será a precisão dos resultados disponibilizados pelos métodos analíticos, fazendo, assim, com que exista uma taxa de falsos positivos mais baixa, enquanto que o processo de tomada de decisão e resposta, a cada incidente, pode ser substancialmente acelerado. É mais importante possuir dados de um número de fontes distintas, do que ter uma única fonte e um grande volume de dados [30]. Segundo Lashkari *et al.* [30], o domínio da cibersegurança obriga a que sejam utilizadas todas as fontes disponíveis, de modo a melhor caraterizar o perfil do utilizador e/ou entidade. Assim, as fontes de dados contribuem para o enriquecimento dos perfis de utilizador ou entidade criados, seja com dados de contexto ou com dados de atividade. Contudo, interessa também identificar quais as que melhor se adequam aos casos de uso e métodos analíticos escolhidos, de forma a otimizar a solução final, sendo sempre importante também conhecer o volume, velocidade e variedade, para que se possa planear a arquitetura devidamente [17].

Os dados fornecidos por estas fontes podem ser primários, sem qualquer tipo de pré-processamento, e gerados em tempo real, como é o caso de *logs* em dispositivos críticos ou pontos de ligação chave, dados de tráfego de rede, ou ainda dados de contexto estruturados ou não estruturados. De seguida, descrevem-se algumas das principais fontes de dados, no contexto de UEBA:

- Diretórios - as informações de diretório presentes na *Active Directory* (AD) ou acessíveis através do *Ligthweight Directory Access Protocol* (LDAP) são uma das fontes mais comuns. Estas permitem compreender a organização, os diferentes papéis, as permissões de acesso, os eventos de autenticação, bem como todos os diferentes atributos que contribuem para a identificação de um utilizador ou entidade;

- Fluxos de tráfego de rede, VPN [2], Proxy [3] e DNS [4] - recolha de dados referentes às comunicações processadas na rede permite aumentar a visibilidade sobre a mesma, possibilitando aceder a informação que inclui volumes de transferência de dados, localizações de transferência, ligações não usuais entre máquinas e ligações não habituais entre fontes internas ou externas. Assim, permitirá detetar atividades de roubo de dados, de comando e controlo, e de movimentação lateral que estejam a ocorrer;
- Dispositivos Terminais - traduzem-se numa fonte rica de informação para a atividade dos utilizadores tanto *online* como *offline*, possibilitando, assim, a obtenção de contexto para ataques baseados em *malware* ou ameaças internas. Esta informação inclui a atividade de manuseamento de ficheiros (por exemplo, imprimir, copiar, colar, publicar e *download*), de utilização da *cloud*, de dispositivos de armazenamento móvel ou de aplicações, etc.;
- Pacotes - uma inspeção aprofundada a este tipo de dados possibilita extrair um conjunto de meta-dados, desde a camada de aplicação até à camada de transporte, com vista a fornecer uma visão mais granular da atividade do utilizador;
- *Logs* - dados enviados por todos os sistemas que possuam capacidade de *logging* (dispositivos terminais, servidores, aplicações, sistemas operativos, bases de dados, correio electrónico, etc.), permitem obter dados como utilizador, endereço IP, grupo data-hora, ação, recurso utilizado, etc.;
- Outras fontes - para além das fontes já mencionadas acima, podem existir outro tipo de fontes que complementam estas. Neste sentido, pode ter-se os seguintes exemplos de fontes: fontes de enriquecimento de dados, fontes de inteligência, alertas de ferramentas de segurança, comportamento na *web* e atividades nas redes sociais, ou ainda dados provenientes dos recursos humanos.

2.5 Ferramentas

A instalação do primeiro SIEM na Altice Portugal já remonta ao ano de 2007, tendo na altura a escolha recaído no *ArcSight* da *Hewlett Packard*. Mais tarde, e considerando o inerente crescimento da infraestrutura da Altice Portugal, foi decidido efetuar uma separação entre a rede interna da organização e as redes dos clientes. Assim, surgiu a necessidade de instalação de um segundo SIEM para monitorização da rede interna, permanecendo o *ArcSight* para monitorização das redes dos clientes. Após realização de um estudo interno, a escolha recaiu sobre o SIEM *Alienvault USM*, o qual foi implementado em 2017 [35].

Já em 2019, e devido a questões relacionadas com o licenciamento da ferramenta *ArcSight*, foi tomada a decisão de substituição dos SIEM em exploração. Com o objetivo de garantir a uniformização entre as aplicações utilizadas pelas empresas do grupo Altice, a escolha de SIEM a utilizar acabou por ser o *IBM QRadar*, uma vez que este já era explorado pela Altice França. Neste sentido, iniciou-se em 2019 o processo de implementação e migração para esta nova ferramenta, ficando a mesma responsável pela monitorização dos eventos tanto da rede interna da Altice Portugal, como das redes dos seus clientes.

Os SIEM, conforme já mencionado neste documento, constituem-se como uma ferramenta vital nos CSOC. Eles são responsáveis pela recolha e agregação de todos os eventos produzidos pelos dispositivos de segurança e, de rede, e pelos diferentes sistemas e aplicações, fazendo a interligação entre estes. Posteriormente, analisa-os e gera alertas para as equipas de segurança. A fonte primária de dados dos

[2] *Virtual Private Network - ligação de comunicações cifrada estabelecida sobre uma rede pública.*

[3] Servidor que funciona como intermediário entre um conjunto de clientes e um conjunto de serviços.

[4] *Domain Name Service - resolução de nomes de domínio em endereços IP e vice-versa.*

SIEM são os *logs*, mas eles têm também capacidade para processar outros tipos de dados, tais como pacotes e fluxos de rede. Os eventos são combinados com informação de contexto, a qual diz respeito a utilizadores, ativos, ameaças e vulnerabilidades. Estes dados podem ser normalizados, permitindo assim a análise a partir de fontes de dados distintas [5]. Segundo Kavanagh *et al.* [5], a definição de SIEM foi necessariamente revista para "[sistema que endereça] *a necessidade dos clientes em analisar eventos em tempo real permitindo uma deteção precoce de ataques direcionados ou roubos de dados, bem como a necessidade de recolha, armazenamento, investigação e produção de relatórios sobre dados dos logs para permitir a resposta a incidentes, análise forense e garantir a conformidade com os regulamentos e normas*".

Conforme ilustrado na figura 2.5, Kavanagh *et al.* [5] mostram o posicionamento relativo entre os dezassete principais competidores no mercado de ferramentas SIEM, dividindo-os em quatro quadrantes (líderes, visionários, desafiadores e nicho) de acordo com as suas caraterísticas. Kavanagh *et al.* [5] colocam metade deles no quadrante de nicho, pois estes fabricantes oferecem soluções mais apropriadas para mercados, ou casos de uso mais específicos. Nenhum recebeu o título de desafiador, categoria esta que procura identificar fabricantes como sendo capazes de provar a execução da sua visão, mas sem possuir um conjunto completo de funcionalidades no seu SIEM. Por outro lado, apenas um recebeu o título de visionário, *LogPoint*, pois disponibiliza uma ferramenta cheia de funcionalidades, mas prova ter uma capacidade limitada na sua execução. A existência de um baixo número de desafiadores e visionários, segundo Kavanagh *et al.* [5], é indicador da existência de um mercado cada vez mais maduro. Por fim, foram nomeados sete fabricantes como pertencendo ao quadrante dos líderes. É, pois, neste cenário que se enquadra a ferramenta SIEM *IBM QRadar*, a qual surge em segundo lugar neste quadrante, apenas ultrapassada de forma pouco significativa pela *Splunk*, quanto à capacidade de execução, mas ficando em primeiro lugar no que toca à completude da visão de eventos de segurança no ambiente onde está a ser explorado.

Por outro lado, os sistemas UEBA estão em tudo relacionados com os SIEM, realizando muitas das mesmas funções, uma vez que eles também colecionam eventos de segurança de toda a organização, os analisam e identificam incidentes de segurança. No entanto, enquanto as soluções UEBA estão essencialmente focadas no lado da análise, os SIEM procuram cobrir o grande volume e variedade de fontes de eventos de segurança, organizando-os para os analistas de segurança, e possibilitando a estruturação de processos nos SOC [36, 37].

No início do seu surgimento, as ferramentas UEBA eram disponibilizadas de forma individual, como uma solução isolada. Contudo, incorporar estas nos SIEM pode trazer muitos benefícios no fortalecimento da segurança das organizações, combinando, assim, a capacidade de integração e informação acedida pelos SIEM com a utilização de métodos e algoritmos analíticos avançados utilizados pelas tecnologias UEBA. Sadowski *et al.* [3] propuseram, em 2018, que as plataformas SIEM deveriam incluir algumas capacidades analíticas avançadas, uma das quais seria as soluções UEBA. Apesar de nunca se referir ao termo de próxima geração de SIEM, a sua visão inclui a disponibilização de uma solução UEBA que possibilite realizar uma análise comportamental de anomalias em eventos de segurança e dados de *logs*. Deste modo, a *Exabeam* [33] considera então estas as seguintes funcionalidades como essenciais nos SIEM modernos:

- Monitorização de utilizadores, incluindo a definição de comportamentos base e analítica avançada para analisar dados de acesso e autenticação, estabelecimento de contexto do utilizador e reportar qualquer comportamento suspeito;
- Métodos analíticos avançados, aplicando estatísticas sofisticadas e modelos quantitativos (como é o caso de *Machine Learning*) nos *logs* de segurança e eventos para detetar anomalias. Estes métodos devem complementar as tradicionais regras e métodos analíticos de correlação existentes nos SIEM tradicionais.

Figura 2.5: Comparação entre os principais Fabricantes de SIEM (extraído de Kavanagh et al. *[5])*

UEBA representa então uma importante e significativa melhoria sobre os sistemas SIEM tradicionais, por um conjunto alargado de motivos. Em primeiro lugar, ultrapassa as limitações das regras de correlação, nomeadamente:

- Falta de contexto - pode impedir a deteção de ataques ou perda de incidentes que nunca tinham sido observados anteriormente, gerando assim falsos negativos;
- Necessidade de muita manutenção;
- Filtragem inapropriada de regras - pode levar a uma resposta a incidentes mais lenta, pois os analistas precisam de filtrar que regras aplicar para determinar que dados são relevantes ou não, num determinado cenário de segurança.

Em segundo lugar, fornece informação de contexto adicional sobre ameaças conhecidas e desconhecidas, identificando-as de forma mais precisa e precoce, bem como funde vários tipos de informação, construindo uma hierarquia de risco por entidades. Desta forma, possibilita poupar tempo aos analistas, aumentando a eficiência dos CSOC, através da redução do número de falsos positivos e aumento da precisão de deteção dos ataques, permitindo também a estes priorizar e focar-se nos alertas que representam um risco mais elevado [38, 39].

Com efeito, nestes últimos anos tem-se já assistido a uma crescente prática de disponibilização das

soluções UEBA lado a lado com os SIEM, numa integração bidirecional, podendo assim ser visto como um complemento perfeito. Este tipo de integração favorece os analistas dos CSOC, os quais podem continuar a fazer uso das suas consolas, ao mesmo tempo que beneficiam também da capacidade de deteção de ataques que as ferramentas UEBA disponibilizam [3].

2.5.1 IBM QRadar SIEM

A *IBM QRadar Security Intelligence Platform* é responsável por fornecer uma arquitetura unificada, a qual integra um sistema para gestão de eventos e informação de segurança (SIEM), gestão de *logs*, deteção de anomalias, análise forense de incidentes e configuração e gestão de vulnerabilidades. O principal sistema desta arquitetura assenta no *IBM QRadar SIEM*, podendo posteriormente ser adicionados outros módulos, tais como *QRadar Risk Manager*, *QRadar Vulnerability Manager* e *QRadar Incident Forensics*, os quais acrescentam funcionalidades específicas. Para além destes módulos, existem também aplicações, como é o caso da *User Behavior Analytics* (UBA), que criam ou estendem funcionalidades no *QRadar*, sendo geralmente desenvolvidas pela comunidade.

O *IBM Security QRadar SIEM*, vulgarmente conhecido apenas por *QRadar*, traduz-se numa arquitetura modular que disponibiliza em tempo-real, visibilidade sobre a infraestrutura das TIC, a qual pode ser utilizada para deteção e priorização de ameaças. É sobre esta arquitetura que é definida a operação do *QRadar*, a qual consiste na utilização de três camadas, conforme ilustrado na figura 2.6. Desta forma, o *QRadar* recolhe, processa, agrega e armazena dados de diferentes fontes existentes na rede, em tempo-real, permitindo-se depois fazer a gestão da segurança através da monitorização, geração de alertas e ofensas, bem como da resposta às diferentes ameaças. Para além do referido, o *QRadar* permite também a realização de pesquisas e produção de relatórios.

Por outro lado, dependendo das necessidades de monitorização, o *QRadar* pode ser disponibilizado em diferentes topologias, isto é, pode consistir numa ferramenta que possui todos os componentes integrados num sistema único ou em várias máquinas que contribuem com papéis específicos, como é o caso dos coletores de eventos, coletores de fluxos, processadores de eventos, processadores de fluxos e nós de dados [6].

Recolha de Dados

A recolha de dados corresponde à primeira camada, onde dados como eventos ou fluxos de tráfego de rede são colecionados a partir da rede da organização. Considerando que a principal função do *QRadar* é garantir a segurança da rede fazendo a monitorização dos eventos e fluxos, esta etapa de recolha de eventos e fluxos reveste-se de grande importância. Para mais, importa também diferenciar os conceitos de evento e fluxo de tráfego de rede, os quais apresentam uma diferença substancial.

Evento representa o registo de uma ação ou acontecimento específico, no ambiente de um utilizador, o qual ocorreu num dado momento temporal bem definido, sendo o referido registo realizado nesse mesmo momento. Como exemplos de eventos tem-se tentativas de autenticação, correio eletrónico, estabelecimento de ligações VPN, acessos a servidores *proxy*, entre outros. O *QRadar* aceita eventos a partir de quaisquer fontes de *logs* que existam na rede, através da utilização de protocolos como *syslog* [5], e *Simple Network Management Protocol* (SNMP) [6]. Para conseguir esta recolha, o *QRadar* recorre a um coletor de eventos, o qual, após os receber, os posiciona numa fila para serem pré-processados. A partir destas filas, os eventos são analisados e normalizados, transformando, assim, os dados dos eventos primários obtidos num formato estruturado, uniforme e utilizável pelo *QRadar*. Posteriormente, os eventos de fontes conhecidas são concatenados, com base nos atributos comuns desses mesmos eventos.

[5] Padrão criado pelo *Internet Engineering Task Force* (IETF) para a transmissão de mensagens de *log* em redes IP.

[6] Protocolo para permitir a gestão e recolha de informação de dispositivos em redes IP.

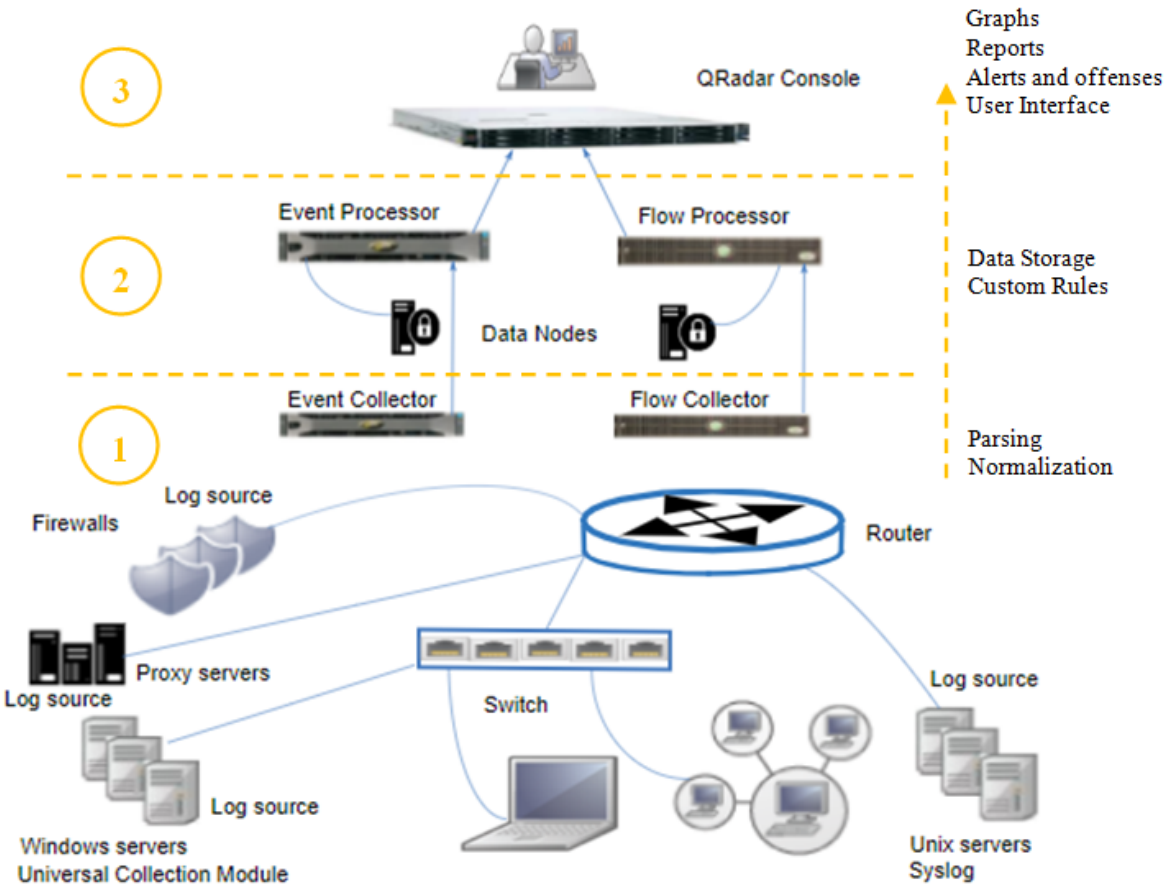

Figura 2.6: Arquitetura do QRadar SIEM (adaptado de IBM [6])

Eventos de fontes novas ou desconhecidas, que não foram detetados no passado, são reencaminhados para o motor de análise de tráfego [40].

Por seu lado, *fluxo de tráfego de rede* representa a informação de atividade de rede ou a informação de uma sessão entre dois *hosts*, a qual pode durar segundos, minutos, horas ou dias. Esta informação (por exemplo, endereços IP, portos, número pacotes, tamanho, entre outros dados) é recolhida pelo colector de fluxos a partir de ferramentas de monitorização de portos (por exemplo, SPAN [7] e TAP [8]), de monitorização de sessões ou a partir de fontes externas de fluxos (por exemplo, *netflow*, *sflow*, *jflow*) [9], traduzindo-a e normalizando-a depois para registos de fluxos passíveis de serem interpretados pelo *QRadar*, os quais vão também para um fila para serem pré-processados. Para tal, considera-se que um fluxo se inicia quando o coletor deteta o primeiro pacote que possui uma identificação única (endereço IP de origem, endereço IP de destino, porto de origem, porto de destino e outras opções específicas de acordo com o protocolo), sendo que, para sessões que abrangem vários minutos, a fila reporta o registo desse fluxo, ao fim de cada minuto, com as métricas atuais. Assim, é possível observar vários registos no *QRadar* com um grupo data-hora igual para o primeiro pacote, mas onde esse grupo data-hora do último pacote vai sendo incrementando ao longo do tempo. Na prática, nenhum pacote é captado integralmente, mas são sim avaliados, ou seja, os contadores respeitantes ao tamanho e número de pacotes vão sendo atualizados para cada registo de fluxo e, no final de cada intervalo de tempo, os contadores são colocados

[7] *Switched Port Analyzer* - também conhecido como *port mirroring*, envia uma cópia de todo o tráfego de rede observado num porto ou até mesmo numa rede, para um outro porto, com o objetivo de ser analisado.

[8] *Terminal Access Point* - dispositivo de *hardware* que permite aceder e monitorizar os eventos numa rede.

[9] Tecnologias distintas, as quais se destinam a recolher o tráfego das redes.

a zero e o registo é enviado para o processador de fluxos. Um fluxo termina quando não for detetada mais nenhuma atividade para esse mesmo fluxo, no período de tempo configurado [40].

Processamento de Dados

A segunda camada corresponde ao processamento dos dados de eventos e ou de fluxos. À semelhança da camada anterior, onde existiam dois coletores distintos para cada um destes tipos de dados, também nesta camada existem dois tipos de processadores diferentes.

O processador de eventos, como o próprio nome indica, processa os eventos recolhidos por um ou mais coletores de eventos. Este processa-os usando um *Custom Rules Engine* (CRE), o qual é responsável por compará-los com regras predefinidas na consola, mantendo um seguimento dos sistemas envolvidos nos incidentes ao longo do tempo, e executando a ação que está definida para a resposta daquela regra despoletada. Este processador de eventos envia também os eventos em tempo-real para a consola, não sendo os mesmos fornecidos a partir de nenhuma base de dados. Contudo, cada processador possui armazenamento local próprio, podendo os eventos ser armazenados no processador, ou num nó de dados específico. Os eventos são, assim, armazenados numa base de dados não relacional, chamada *Ariel*, a qual é atualizada minuto a minuto, após estes serem processados. A existência dos nós de dados permite aumentar o espaço de armazenamento, bem como melhorar o desempenho das pesquisas através do aumento da velocidade destas [40].

Quanto aos processadores de fluxos, estes são responsáveis pelo processamento dos fluxos recebidos de um ou mais coletores de fluxos, bem como de eventuais fontes externas, conforme já mencionado acima. De resto, apresentam um comportamento em tudo semelhante ao descrito anteriormente para os processadores de eventos [40].

Pesquisa de Dados

A última camada situada no topo desta arquitetura corresponde à consola de interface com o utilizador. Deste modo, é através desta camada que os dados ficam disponíveis aos utilizadores para que estes executem as pesquisas, a monitorização de eventos e fluxos em tempo-real, análises, relatórios, investigações a alertas e ofensas, visualização de informação dos ativos, bem como diferentes tarefas de administração [6].

Quando um dos processadores de eventos ou fluxos envia uma notificação para a consola do *QRadar*, fá-lo, nomeadamente, para o seu componente *Magistrate*. Esta notificação significa que uma dada regra foi despoletada, sendo automaticamente gerada uma ofensa, e podendo também ser executadas outras ações ou respostas, como notificações, envio de correio eletrónico, entre outros. As ofensas são criadas e geridas por este componente, o qual também as guarda numa base de dados dedicada, residente na consola [40].

Por fim, conforme referido no início desta secção, podem ser adicionadas aplicações ao *QRadar SIEM*. Por omissão, elas são instaladas na consola do *QRadar*. No entanto, por questões de desempenho, as mesmas devem ser instaladas numa máquina dedicada, tornando possível a disponibilização de mais memória, armazenamento e capacidade de processamento, sem afetar o desempenho e capacidade de processamento da consola. Aplicações, como é o caso da *User Behavior Analytics* (UBA), requerem mais recursos dos que os disponíveis, presentemente, na consola.

2.5.2 IBM QRadar UBA App

A aplicação *User Behavior Analytics* (UBA) é uma ferramenta que possibilita a deteção de ameaças internas numa organização. Ela constitui-se como uma extensão ao *IBM QRadar SIEM*, e é um componente opcional da *IBM QRadar Security Intelligence Platform*. Esta aplicação é disponibilizada de modo

gratuito, e pode ser descarregada a partir do repositório *IMB Security App Exchange*. Esta ferramenta disponibiliza, de forma pronta a utilizar, um conjunto de regras e algoritmos que se ligam diretamente ao motor analítico avançado do *QRadar*. Esta solução integrada é desenhada para adicionar contexto de utilizadores a *logs*, tráfego de rede, vulnerabilidades e outros dados recolhidos pelo *QRadar SIEM*. Usando estes dados anteriores, estabelece a linha base dos padrões de acesso e atividades normais dos utilizadores, para identificar categoricamente os comportamentos desviantes, gerar valores de risco para os utilizadores e fornecer aos analistas uma visão dos utilizadores potencialmente comprometidos ou com um risco mais elevado associado. Como resultado, é criado e monitorizado um perfil de risco por utilizador, fazendo uso da lógica desta aplicação, bem como de diferentes algoritmos de *machine learning*. Deste modo, o principal objetivo da aplicação UBA passa pela deteção de comportamentos anómalos e maliciosos de utilizadores, baseando-se num resultado de risco por utilizador, o qual é designado por *risk score*. A aplicação UBA permite então:

- Estabelecer automaticamente uma linha base das atividades normais dos utilizadores;
- Identificar comportamentos anómalos de utilizadores, os quais podem indicar uma ameaça interna ativa;
- Determinar perfis de risco dos utilizadores, priorizando os utilizadores e as atividades de risco mais elevado na organização;
- Alertar os analistas de segurança para potenciais ataques internos.

Por outro lado, com a instalação desta aplicação, as equipas de segurança também saem beneficiadas, tornando o seu trabalho mais produtivo e eficiente, ajudando-as nomeadamente a [41]:

- Identificar os utilizadores de alto risco antes de se tornarem desonestos;
- Expor os ciber criminosos a operar sob credenciais comprometidas;
- Monitorizar e gerir as atividades dos utilizadores com privilégios especiais;
- Monitorizar o acesso aos principais ativos da organização;
- Compreender e quantificar o perfil de risco do ambiente;
- Detalhar e analisar de forma integrada as ameaças potenciais, ganhando também conhecimento sobre a ação corretiva.

As duas principais contribuições da aplicação UBA para o *QRadar* são as seguintes funções [7]:

- Perfil Risco - criação de um perfil de risco por utilizador, construído à custa da atribuição de valores de risco a diferentes casos de uso de segurança. Os exemplos podem incluir desde simples regras e verificações, tais como acesso a *websites* maliciosos, ou métodos analíticos mais avançados que fazem uso de *machine learning*. O valor de risco é atribuído a cada um dependendo da severidade e confiabilidade do incidente detetado.
- Unificação de Identidades de Utilizadores - conseguido através da procura em contas díspares por um dado utilizador, isto é, devido à importação de dados da AD, LDAP ou ficheiro CSV [10] a aplicação UBA pode ser ensinada para saber que contas pertencem ou traduzem a identidade de um utilizador, permitindo assim combinar o risco e os diferentes eventos de segurança através das diferentes identificações desse utilizador nesses mesmos eventos.

[10] *Comma-separated values* - ficheiro de texto delimitado que utiliza vírgulas para fazer a separação entre os valores

Como Funciona

A aplicação UBA está construída sobre a *framework* da aplicação *QRadar*. A figura 2.7 ilustra a integração e a forma como o *QRadar SIEM* e a aplicação UBA se interligam. Ela trabalha, assim, de forma conjunta com este, para recolher dados dos utilizadores que se encontram dentro da rede da organização [7]:

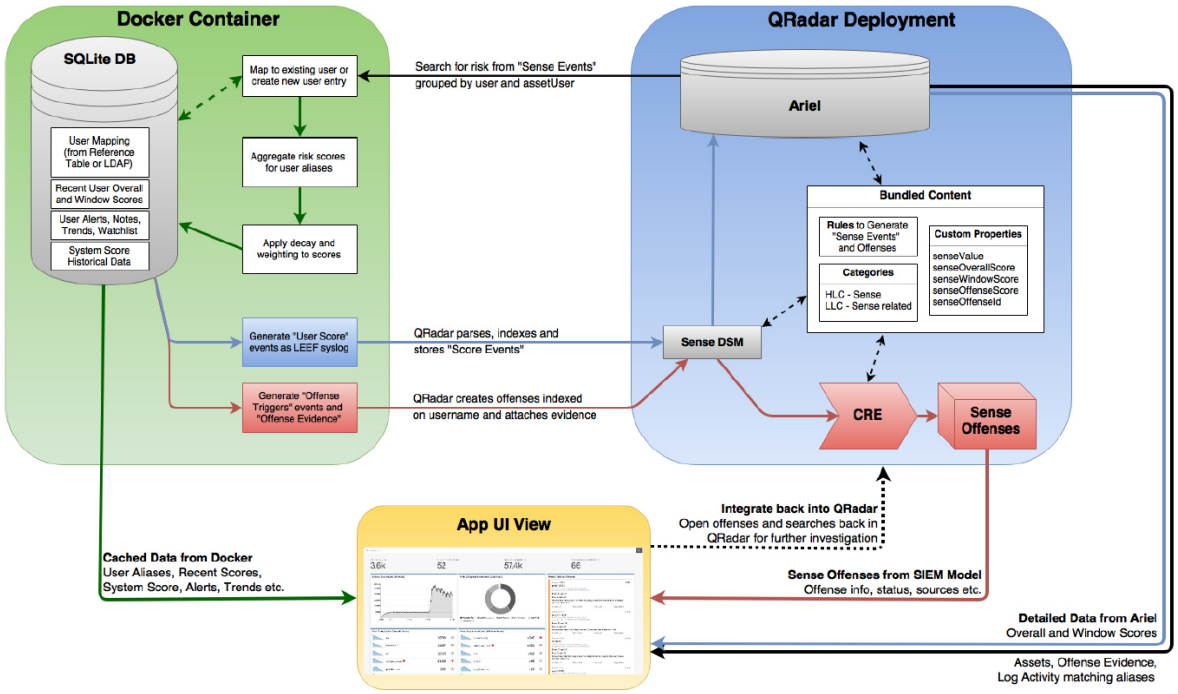

Figura 2.7: Funcionamento aplicação da UBA (extraído de IBM[7])

1. O *QRadar SIEM* recebe dados originários de fontes como *logs* e tráfego de rede, os quais são armazenados numa base de dados chamada *Ariel*. Regra geral, a aplicação UBA requer a existência de fontes de dados que forneçam um nome de utilizador. Contudo, quando não existe qualquer nome, e caso a pesquisa por ativos esteja ativada, então ela tenta procurar pelo utilizador nessa tabela de ativos. Se não se conseguir determinar nenhum utilizador, então a aplicação UBA não processa o evento;
2. A aplicação UBA especifica um conjunto de regras, as quais mapeiam um conjunto de casos de uso, definidos a montante. Estas regras são em tudo semelhantes às regras de correlação do *QRadar SIEM*, estando especificamente afinadas para esta. Elas procuram, assim, por determinadas anomalias e, caso se encontrem ativas, geram novos eventos, designados por *Sense Event*, os quais são lidos e interpretados pela aplicação UBA. Por omissão, esta já disponibiliza um grande conjunto de regras, oferecendo também a possibilidade de alteração ou adição de novas regras, de acordo com as necessidades específicas de cada cliente. Para além destas regras, a aplicação disponibiliza ainda métodos analíticos baseados em *machine learning*, sendo para tal necessário realizar a instalação da aplicação *Machine Learning*. Esta funciona de forma semelhante às regras, sendo o seu funcionamento explicado de forma mais detalhada abaixo;
3. Cada regra e modelo analítico possui um valor de risco associado, designado por *senseValue*, o qual indica a severidade da anomalia encontrada. Cada vez que a atividade de um utilizador, materializada num conjunto de ações, despoleta uma determinada regra ou modelo analítico, este valor é adicionado ao resultado de risco do utilizador em questão. Quanto maior for o

número de violações realizadas por um utilizador, maior será o seu resultado, originando assim um perfil de risco mais elevado. Na prática, a aplicação UBA retira o *senseValue* e o nome de utilizador do *Sense Event* e aumenta o *Risk Score* desse mesmo utilizador com um valor igual ao contido no *senseValue*;

4. O resultado de risco de um utilizador é a agregação de todos os eventos de risco detetados pelas regras ou modelos analíticos. Quanto maior for este valor, maior será a probabilidade de esse utilizador confiável da organização representar um risco de segurança, devendo a sua atividade ser investigada de forma aprofundada. Contudo, caso não ocorra nenhum novo evento, o resultado do risco de cada utilizador vai diminuindo ao longo do tempo. O parâmetro designado por *decay* controla o fator de redução do valor de risco por hora;
5. Quando o resultado de risco de um utilizador excede um limite bem definido, o qual pode ser devidamente parametrizado nesta aplicação, esta envia um evento que despoleta a regra "*UBA : Create Offense*", sendo criada uma ofensa para este utilizador, permitindo assim a geração de um alerta, para análise, por um analista de segurança;

Importação de Dados dos Utilizadores

Conforme já mencionado atrás, uma das principais funções da aplicação UBA é a unificação de identidades de utilizadores. Por outro lado, caso não se consiga atribuir um evento a um utilizador, esse evento não será processado. Desta forma, esta etapa de aquisição de dados referentes aos diferentes utilizadores reveste-se de uma importância crucial para o sucesso de utilização desta ferramenta.

Utilizando a AD ou o LDAP, um único utilizador pode ser descrito com múltiplos atributos usados em diferentes fontes de dados. Assim, estas contribuem necessariamente para a identificação da atividade de um único utilizador, através das várias fontes, uma vez que um mesmo utilizador pode estar presente nessas diferentes fontes, com diferentes identificações. A este processo de agregar todas as atividades de um utilizador, realizadas nos diferentes sistemas, num único indivíduo, dá-se o nome de *coalescing* (concatenação).

Neste sentido, e agregado à aplicação UBA encontra-se a aplicação *Reference Data Import*, a qual possibilita obter informação de identificação contextual a partir de múltiplas fontes LDAP. Esta aplicação é instalada de forma automática quando é realizada a instalação da aplicação UBA, e o seu principal objetivo é permitir a importação de dados e/ou atributos de utilizadores da organização a partir da AD, LDAP ou ficheiro CSV para uma tabela de referência. Esta tabela é depois utilizada pela aplicação UBA, podendo também ser usada pelas pesquisas e regras do *QRadar*.

Machine Learning

A aplicação UBA, para além das regras, possui também um conjunto de métodos analíticos comportamentais avançados e algoritmos de aprendizagem automática, os quais contribuem para o aumento da robustez do motor analítico do *QRadar SIEM*. Estes estão todos integrados e são disponibilizados através da aplicação *Machine Learning* (ML), a qual é fornecida conjuntamente com a aplicação UBA. A aplicação ML constitui-se, assim, como uma ferramenta adicional que estende as capacidades da aplicação UBA, adicionando casos de uso mais avançados e detalhados para os algoritmos de ML, permitindo construir perfis e executar agrupamentos numa base temporal. Ela é instalada a partir da aplicação UBA, sendo a mesma opcional [41, 7].

A aplicação ML recorre a cinco algoritmos de aprendizagem automática distintos, os quais permitem analisar as atividades dos utilizadores, e criar um modelo comportamental para cada um deles, de forma individual. A aplicação processa os dados provenientes das fontes configuradas no *QRadar SIEM*,

num período precedente de quatro a seis semanas e, poucas horas após a sua instalação e configuração, ela é capaz de compreender os padrões normais de atividade de cada utilizador monitorizado. Posteriormente, estes algoritmos são capazes de prever as atividades futuras dos utilizadores e a sua frequência. Assim, quando estes saírem daquilo que é o padrão previsto, os algoritmos sinalizam estas atividades como comportamento anómalo. De forma análoga ao que acontece com as regras, quando ocorre esta sinalização, a aplicação adiciona o *senseValue* do *Sense Event* gerado ao *Risk Score* desse utilizador. Mais, esta aplicação permite também aos analistas de segurança ativar a opção de escalar o *senseValue* de forma dinâmica, dependendo da magnitude do desvio que este evento originou relativamente ao padrão normal. Assim, o *senseValue* resultante é calculado atráves da multiplicação do valor de risco configurado no modelo analítico pelo fator baseado no quão distante é esse desvio. Este fator recebe o nome de *senseScore*. Deste modo, se um utilizador apresentar um fator de desvio de cinco pontos em relação à média, e for utilizado o valor de risco por omissão de cinco pontos, o utilizador verá adicionados ao seu *Risk Score* um valor de vinte e cinco pontos. Apesar das vantagens que esta aplicação acrescenta à plataforma *QRadar*, devido às necessidades de processamento dos seus algoritmos, ela faz um uso intensivo dos recursos da máquina onde está instalada e configurada [41, 7].

Deste modo, os algoritmos de aprendizagem automática na aplicação UBA podem ser utilizados em três âmbitos distintos [41]:

- Deteção de desvios comportamentais dos utilizadores em relação a si próprios - fazendo uso de uma análise *Multimodal Gaussian* [11], monitoriza os utilizadores através de um conjunto de diferentes categorias de alto nível de eventos, tais como autenticação, comportamento na rede, atividade de aplicações, *malware* ou outra atividade de *software* malicioso, entre outros. As anomalias detetadas são mostradas em gráficos de aranha, podendo ser monitorizadas as seguintes: aumento anormal na atividade do utilizador ao longo do tempo, desvio num tipo específico de atividade do utilizador (por exemplo, pedidos de autenticação), desvio na postura de risco do utilizador, aumento anormal da taxa de execução de atividades com maior risco associado, desvio ou aumento na atividade no sentido local-remoto, e alterações realizadas ao sistema do utilizador (por exemplo, novas instalações de *software*);
- Deteção de alterações na atividade do utilizador *versus* frequência - cria um modelo de distribuição de atividade e frequência, ao longo do tempo, para cada utilizador, com recurso a modelos como *Latent Dirichlet Allocation* (LDA) [12] e *Kullback-Leibler Divergence* (KLD) [13]. Estes modelos podem ajudar a detetar situações em que há um comprometimento das credenciais de um utilizador, ou quando um utilizador com credenciais válidas altera a frequência com que habitualmente executa determinadas tarefas;
- Deteção de desvios anómalos em relação aos grupos de pares - este terceiro âmbito faz uso dos modelos de comportamento individuais criados anteriormente, e constrói agrupamentos de utilizadores com comportamentos similares. Para tal, recorre a algoritmos como *Gaussian Mixture* [14] e *Jaccard Similarity* [15] para identificar e agrupar os utilizadores. De seguida, utiliza a *Kullback-Leibler Divergence* para detetar quando um dado utilizador se desvia do seu grupo. Sempre que se desvia, o *risk score* é aumentado, conforme já explicado anteriormente. Assim,

[11] Distribuição de probabilidades contínua com dois ou mais modos.

[12] Modelo estatístico generativo que permite que conjuntos de observações sejam inferidos por grupos não observados que explicam porque algumas caraterísticas são similares.

[13] Medida de como uma distribuição de probabilidades diverge de uma segunda distribuição de probabilidades (distância).

[14] Modelo probabilístico que assume que todos os pontos são gerados a partir de uma mistura finita de distribuições Gaussianas com parâmetros desconhecidos, para representar a presença de sub-populações numa população geral, sem requerer que um conjunto de dados observado identifique a sub-população à qual uma observação individual pertence.

[15] Coeficiente que permite medir a similaridade entre conjuntos com amostras finitas.

com este tipo de abordagem de análise de comportamento de grupos de utilizadores, torna-se possível a existência de uma outra perspetiva sobre as atividades do utilizador, ajudando, desta forma, a identificar qualquer atividade anómala ou maliciosa, quando este se desvia do grupo de utilizadores com papéis e responsabilidades similares.

Para finalizar, verifica-se que são disponibilizados, por omissão, um conjunto de 11 modelos de casos de uso, tais como atividade de acesso, distribuição de atividade, atividade de autenticação, atividade agregada, atividade suspeita, exfiltração de dados, entre outros. Estes procuram cobrir, de forma genérica (alto nível), as principais necessidades de monitorização de comportamento de segurança de utilizadores, de forma transversal às organizações. Os modelos podem então ser ativados e/ou podem ser editados os seus parâmetros originais. Adicionalmente, podem também ser criados novos modelos de ML próprios, fazendo uso, ou não, de *templates* incluídos nesta aplicação, por forma a suportar casos de uso específicos. De notar que, apesar de não existir um número limite para os modelos existentes/configurados, apenas podem estar ativos, num dado momento, 17 modelos. Todas estas funcionalidades podem ser encontradas a partir da versão 3.3.0 desta aplicação.

Capítulo 3

Análise e Desenho da Solução

3.1 Definição e Análise do Problema

O problema exposto pela Altice Portugal para a realização deste projeto assenta na atual necessidade de monitorização do comportamento normal de utilizadores confiáveis da organização, visando futuramente, detetar e analisar anomalias comportamentais (comportamentos desviantes) e, assim, identificar eventuais utilizadores maliciosos ou mal intencionados.

Em primeiro lugar, para melhor abordar este problema, devem ser definidas e caraterizadas as necessidades de monitorização específicas. Neste sentido, importa começar por identificar e priorizar uma lista de casos de uso, os quais refletem as principais preocupações de monitorização ao nível das possíveis ameaças internas. Em segundo lugar, e após esta identificação dos casos de uso mais prioritários, interessa compreender e identificar quais os dados de atividade ou contexto, e respetivas fontes, que contribuem para uma caraterização mais precisa dos casos de uso anteriores. Para endereçar este problema, deverá ser explorada e utilizada a aplicação UBA, a qual se constitui como parte integrante da plataforma *IBM QRadar*, utilizada pelo CSOC desta organização.

Deste modo, nesta fase de análise inicial do problema, existem algumas perguntas obrigatórias que deverão ser feitas para que possam ser identificados a nível macro os principais casos de uso. Esta fase constitui-se como fundamental, uma vez que o sucesso das etapas seguintes, de escolha das fontes de dados e, subsequente implementação, depende fortemente desta fase. Como exemplos dessas perguntas tem-se:

1. Qual o tipo de trabalhadores/empregados que pretendem monitorizar?
2. Quais os tipos de ameaças internas que pretendem detetar?
3. Quais os tipos de atividades de interesse que pretendem monitorizar como possíveis indicadores de comportamento malicioso ou anómalo?
4. Qual a granularidade temporal pretendida para a respetiva monitorização?

De forma a responder a estas questões, deverão ser consideradas as políticas de segurança e o apetite ao risco da organização, bem como a visão estratégica do *Chief Information Security Officer* (CISO) da Altice Portugal, a fim de garantir que a estratégia de monitorização e gestão corresponde às principais necessidades desta. Nesta sequência, deve considerar-se as seguintes respostas referentes às perguntas acima, respetivamente:

1. Pretende-se monitorizar todos os utilizadores que se encontrem ligados à rede da Altice Portugal, incluindo, assim, neste âmbito utilizadores com privilégios normais, com privilégios especiais e até mesmo utilizadores temporários ou associados a prestadores de serviços;

2. A principal ameaça a monitorizar consiste nos utilizadores maliciosos ou mal intencionados;
3. Os principais tipos de atividade de interesse a monitorizar são o abuso de credenciais e a existência de discrepâncias tempo-espaciais;
4. A granularidade temporal desejada para a monitorização dos utilizadores identificados anteriormente é aquela que permita uma deteção em tempo quase real.

Assim, a partir dos pressupostos acima, os casos de uso a nível macro que serão considerados no âmbito deste projeto são:

- Tentativa de Ataque de Força Bruta;
- Utilização de Estação de Trabalho Não Habitual;
- Utilização de Estação de Trabalho em Períodos Não Habituais.

Por outro lado, importa também considerar, na fase de implementação dos casos de uso acima identificados, a adoção de uma estratégia que vise minimizar a ocorrência de falsos positivos. Para tal, deverá ser analisada a atividade normal de *logs* e de rede, por forma a tornar possível identificar e filtrar os conjuntos de dados que não contribuam para o correto treino dos modelos de *Machine Learning*.

3.2 Definição das Fontes de Dados

A definição das fontes de dados e respetivos eventos de segurança é uma etapa considerada crítica e essencial no processo de definição dos casos de uso. Tal sucede pois interessa definir e escolher apenas os eventos e as fontes relevantes, que melhor contribuem para a caraterização de cada caso de uso. Apesar de a arquitetura atualmente implementada pela Altice Portugal para a recolha e envio de *logs* para o sistema SIEM em exploração não fazer parte do âmbito deste trabalho, importa compreender a mesma, de modo a permitir a definição e escolha de fontes e eventos exequíveis.

Conforme referido anteriormente, o *QRadar* está preparado para recolher dados originados de eventos ou fluxos. No entanto, à data de realização deste projeto, a recolha de fluxos não era exequível, pelo que os dados provenientes de fontes com esta origem não serão considerados, uma vez que não é possível a sua utilização. Por outro lado, importa dizer também que a recolha de eventos, conforme ilustrado na figura 3.1, diz respeito apenas aos *logs* de segurança dos sistemas *Microsoft Windows*, mais concretamente das estações de trabalho.

A figura 3.1 demonstra o processo de recolha, gestão e centralização de eventos de segurança de interesse, das estações de trabalho *Microsoft Windows* na Altice Portugal. Recorrendo à utilização do *Windows Event Collection*, os eventos de segurança são encaminhados das fontes de eventos (*Windows Event Forwarding* - WEF) para os diferentes coletores de eventos (*Windows Event Collector* - WEC), onde são armazenados. Os sistemas operativos da *Microsoft* permitem, assim, através da criação de subscrições, configurar o reencaminhamento de uma cópia dos seus eventos para um repositório centralizado (WEC) onde estes podem ser visualizados, tratados e analisados em conjunto, ou seja, os coletores de eventos são utilizados como repositório centralizado de todos os eventos gerados pelas estações de trabalho a monitorizar. Mais, eles permitem a criação das tais subscrições, neste caso iniciadas pela fonte, podendo, desta forma, através de políticas de grupo (*Group Policy* - GPO), especificar que fontes e eventos são recolhidos e em que registo são armazenados no coletor.

Para permitir a gestão e monitorização do ambiente *Windows Event Collection*, vulgo vários coletores (WEC), a Altice Portugal recorre à ferramenta *Supercharger* desenvolvida pela *Logbinder*. Esta, como se observa, apresenta um servidor designado de *Supercharger Manager*, responsável por gerir todos os coletores de forma centralizada e monitorizar cada objeto. Depois, fazendo uso de agentes designados de *Supercharger Agent* ou *Supercharger Controller* instalados em cada um dos coletores de

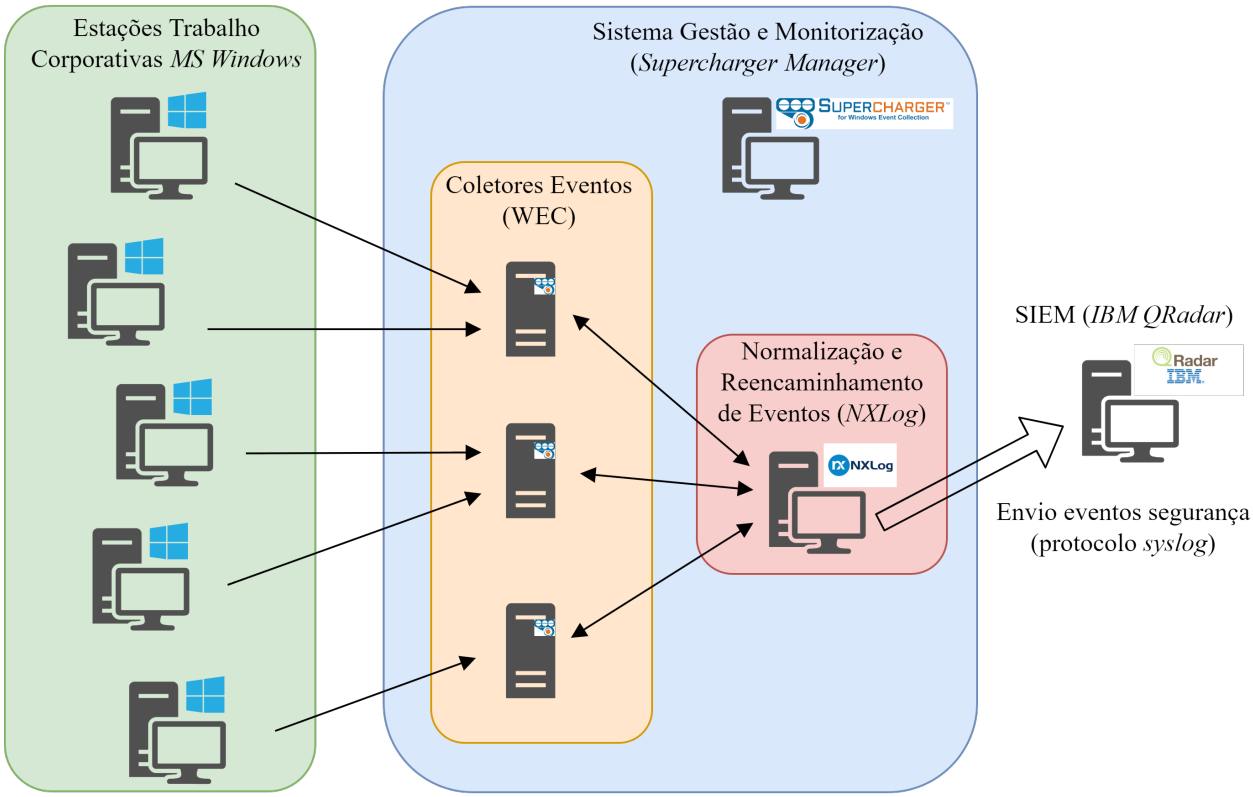

Figura 3.1: Arquitetura para a Recolha de Eventos

eventos, consegue obter os dados de cada coletor, bem como processar automaticamente os comandos enviados pelo *Supercharger Manager*.

Posteriormente, por forma a filtrar, normalizar e reencaminhar os eventos recebidos nos coletores (WEC) para o SIEM, é utilizada a aplicação *NXLog Community Edition*. Assim, apenas os eventos que serão utilizados por este são encaminhados, através do protocolo *syslog*.

Após esta visão geral sobre a arquitetura implementada na Altice Portugal para a recolha de eventos de segurança, o próximo passo na definição das fontes de eventos traduz-se na identificação das caraterísticas que contribuem para a definição de um dado caso de uso. Assim, para cada um destes, as principais caraterísticas são:

1. Tentativa de Ataque de Força Bruta:
    - Nome de utilizador;
    - Evento de falha na autenticação, seja por utilizador ou palavra-passe inválidos.
2. Utilização de Estação de Trabalho Não Habitual:
    - Nome de utilizador;
    - Nome da estação de trabalho;
    - Evento de autenticação bem sucedido numa máquina;
    - Evento de falha na autenticação ou negação de acesso.
3. Utilização de Estação de Trabalho em Períodos Não Habituais:
    - Nome de utilizador;
    - Nome da estação de trabalho;

- Evento de autenticação bem sucedido numa máquina;
- Evento de falha na autenticação ou negação de acesso;
- Grupo data-hora de acessos com autenticação e autorização bem sucedida.

A quantidade de eventos de segurança distintos disponibilizados pelo sistema operativo *Windows*, apesar de variar de versão para versão, traduz-se num número demasiado elevado. Nesta sequência, por questões de desempenho e escalabilidade, torna-se necessário definir quais os eventos de interesse que devem ser recolhidos e armazenados. Para realizar a correta análise e seleção dos eventos necessários recorreu-se à informação oficial disponibilizada pela *Microsoft*, bem como à existente na página *web* da *Ultimate Windows Security*[16].

3.3 Definição dos Casos de Uso

Conforme revisão do estado da arte, os casos de uso traduzem uma condição ou evento específico, geralmente relacionados com uma ameaça concreta, isto é, os casos de uso devem refletir as atividades que se pretende detetar e/ou monitorizar [3]. Considerando os três principais casos de uso, bem como as caraterísticas que contribuem para a sua definição, torna-se agora possível especificar e detalhar os mesmos. Neste sentido, apresenta-se abaixo a definição dos casos de uso:

Tentativa de Ataque de Força Bruta

* Objetivo: este caso de uso endereça o objetivo de detetar, monitorizar e responder a tentativas de abuso de credenciais. Este abuso de credenciais compreende tentativas de exploração de palavras-passe conhecendo nomes de utilizador, ou tentativas de exploração de ambos;
* Ameaça: o abuso de credenciais, em caso de sucesso, torna possível o acesso a sistemas por pessoas não autorizadas. A partir daqui, e recorrendo a ataques de movimentação lateral, podem ser comprometidas outras contas de utilizador e, consequentemente, a confidencialidade, integridade e disponibilidade da informação e sistemas;
* *Stakeholders*: CISO, Analistas do CSOC;
* Atores: todos os utilizadores da Altice Portugal;
* Prioridade: baixa;
* Parâmetros de Entrada: intervalo de confiança compreendido entre 95% e 99% para despoletar anomalias;
* Parâmetros de Saída: a ocorrência deste evento contribui com um valor de risco de 10 pontos;
* Fontes de Dados:
  - Alto Nível: estações de trabalho
  - Baixo Nível: sistema operativo *Windows* 10, 7, Vista, com os seguintes identificadores (ID) de eventos:
    # Falha na autenticação de uma conta - 4625 (considerar os seguintes *Sub Status Code*: 0xC0000064 - nome de utilizador não existe; 0xC000006A - nome de utilizador correto, mas palavra-passe inválida; 0xC000006D - nome de utilizador incorreto ou informação de autenticação errada; 0xC000006E - nome utilizador desconhecido ou palavra-passe inválida).

[16] Divisão da *Monterey Technology Group*, responsável também pela *Logbinder - Supercharger*.

Utilização de Estação de Trabalho Não Habitual

* Objetivo: este caso de uso endereça o objetivo de detetar o acesso ou tentativa, e utilização de uma dada estação de trabalho, com sucesso, ou não, por um utilizador que nunca acedeu à mesma, de acordo com o seu comportamento e atividade passada;
* Ameaça: este caso de uso reflete também a preocupação referente ao abuso de credenciais. Neste caso, um dado utilizador pode fazer uso das suas credenciais, autenticar-se em estações de trabalho para as quais tem permissões de acesso mas, considerando o desempenho das suas tarefas normais atribuídas, não tem necessidade de aceder. Tal pode contribuir para a quebra da confidencialidade e integridade dos sistemas e informação;
* *Stakeholders*: CISO, Analistas do CSOC;
* Atores: todos os utilizadores da Altice Portugal;
* Prioridade: média;
* Parâmetros de Entrada: intervalo de confiança compreendido entre 95% e 99%para despoletar anomalias;
* Parâmetros de Saída: a ocorrência deste evento contribui com um valor de risco de 20 pontos;
* Fontes de Dados:
  - Alto Nível: estações de trabalho
  - Baixo Nível: sistema operativo *Windows* 10, 7, Vista, com os seguintes identificadores (ID) de eventos:
    # Autenticação de uma conta com sucesso - 4624 (considerar os seguintes *Logon Type*: 2 - um utilizador autenticou-se neste computador; 7 - este computador foi desbloqueado; 10 - um utilizador autenticou-se neste computador de forma remota ; 11 - um utilizador autenticou-se neste computador fazendo uso de credenciais de rede armazenadas localmente nesta mesma máquina);
    # Falha na autenticação de uma conta - 4625 (considerar os seguintes *Sub Status Code*: 0xC0000070 - autenticação a partir de estação de trabalho não autorizada; 0xC0000072 - autenticação com conta desativada; 0xC000015B - utilizador sem permissões para o tipo de autenticação pedido; 0xC0000193 - conta de utilizador expirada);
    # Tentativa de autenticação fazendo uso de credenciais de forma explícita - 4648;
    # Negação de acesso ao ambiente de trabalho remoto a um utilizador - 4825.

Utilização de Estação de Trabalho em Períodos Não Habituais

* Objetivo: este caso de uso estabelece o objetivo de detetar o acesso ou tentativa e utilização de uma dada estação de trabalho por um determinado utilizador, num período identificado como não habitual, considerando o seu comportamento e atividade passada;
* Ameaça: este caso de uso endereça simultaneamente a preocupação referente ao abuso de credenciais e à existência de discrepâncias temporais. Neste caso, as credenciais de um dado utilizador podem ser comprometidas, por exemplo através de ataques de *phishing* e, posteriormente, utilizadas para se autenticar em estações de trabalho. Desta forma, caso este acesso ocorra num período temporal distinto do praticado habitualmente pelo verdadeiro utilizador,

este acesso ilegítimo será detetado. Este caso de uso visa contribuir para o evitar da quebra da confidencialidade e integridade dos sistemas e informação;

* *Stakeholders*: CISO, Analistas do CSOC;
* Atores: todos os utilizadores da Altice Portugal;
* Prioridade: alta;
* Parâmetros de Entrada: intervalo de confiança compreendido entre 95% e 99% para despoletar anomalias;
* Parâmetros de Saída: a ocorrência deste evento contribui com um valor de risco de 30 pontos;
* Fontes de Dados:
  - Alto Nível: estações de trabalho
  - Baixo Nível: sistema operativo *Windows* 10, 7, Vista, com os seguintes identificadores (ID) de eventos:
    # Autenticação de uma conta com sucesso - 4624 (considerar os seguintes *Logon Type*: 2 - um utilizador autenticou-se neste computador; 7 - este computador foi desbloqueado; 10 - um utilizador autenticou-se neste computador de forma remota ; 11 - um utilizador autenticou-se neste computador fazendo uso de credenciais de rede armazenadas localmente nesta mesma máquina);
    # Falha na autenticação de uma conta - 4625 (considerar os seguintes *Sub Status Code*: 0xC0000070 - autenticação a partir de estação de trabalho não autorizada; 0xC0000072 - autenticação com conta desativada; 0xC000015B - utilizador sem permissões para o tipo de autenticação pedido; 0xC0000193 - conta de utilizador expirada);
    # Tentativa de autenticação fazendo uso de credenciais de forma explícita - 4648;
    # Negação de acesso ao ambiente de trabalho remoto a um utilizador - 4825.

Capítulo 4

Implementação e Testes

4.1 Instalação e Configuração da Aplicação UBA

Previamente à implementação dos casos de uso detalhados no capítulo anterior, importa começar por descrever o ambiente de qualidade a utilizar, bem como as condições que devem estar reunidas para possibilitar uma implementação. Estas materializam-se ao nível da instalação e configuração das aplicações UBA e ML da *IBM*, as quais serão aprofundadas seguidamente.

O ambiente de qualidade disponibilizado pela organização Altice Portugal consiste numa estação virtualizada com o sistema operativo *CentOS Linux release 7.7.1908 (Core)*, com quatro processadores virtuais *Intel(R) Xeon(R) CPU E5-2630v4 2.20GHz*, 28 GB de memória RAM e 300 GB de armazenamento em disco. Sobre esta plataforma foi instalado o *IBM QRadar Security Intelligence - Community Edition*, na sua versão 7.3.1 *Build* 20180723171558, atualizado (*patched*) para a versão 7.3.1 *Build* 20190228154648. Conforme mencionado no estado da arte, o *QRadar* tem capacidade para recolher dados, os quais representam eventos ou fluxos de tráfego de rede. Neste sentido, para este projeto foram disponibilizados os eventos identificados no capítulo 3, recolhidos por seis servidores *Supercharger* e respeitantes aos *logs* de segurança dos sistemas *Microsoft Windows*, apenas referentes às estações de trabalho. Importa dizer ainda que todos estes eventos recolhidos constituem-se como uma réplica dos que são enviados paralelamente para o ambiente de produção, pelo que o ambiente de qualidade processará dados reais da organização.

Considerando os objetivos deste projeto, a primeira etapa consistiu na instalação da aplicação UBA. A instalação foi executada, conforme indicado no Anexo A, tendo sido instalada a versão 3.4.0, a qual é compatível com a versão 7.3.1 do *QRadar*. Esta é disponibilizada pela *IBM* sem qualquer custo adicional. Posteriormente, a etapa seguinte traduziu-se na configuração desta mesma aplicação, a qual seguiu os passos e ações discriminados no Anexo B. Assim, destacam-se as principais configurações realizadas:

- Configuração do *Authorization Token* - essencial para visualizar a informação na aplicação UBA;
- Configuração de *Content Package Settings* - consiste na ativação de um pacote de regras, das propriedades personalizadas e dos conjuntos de dados de referência utilizados por esta aplicação, com valores disponibilizados por omissão. Caso esta opção não seja ativada, terão de ser configuradas regras próprias;
- Configuração de *Application Settings* - consiste num conjunto de parâmetros que permitem definir o comportamento da aplicação. Deste modo, foram alterados dois parâmetros, *Risk Treshold* e *Decay Risk by this factor per hour*, tendo nos restantes permanecido o valor por omissão. No primeiro caso definiu-se o *Risk Treshold* como estático e com um valor 70, ou

seja, sempre que o resultado de risco de um utilizador exceder este valor, será gerada uma ofensa. Este valor foi escolhido pois qualquer dos casos de uso a implementar, por si só, não é suficiente para gerar uma ofensa. Quanto ao segundo parâmetro, definiu-se como fator de amortecimento o valor de 0.05, de forma a obter uma baixa redução do resultado de risco, permitindo, assim, identificar a atividade dos utilizadores de uma forma mais granular, bem como tornar mais estável (sem alterações constantes) a lista de utilizadores com um resultado de risco mais elevado. Esta ação visa favorecer um melhor estudo e análise neste projeto;

- Ativação de *Indexes* - conforme referido no manual de utilizador [42], e sendo um dos objetivos deste projeto garantir um bom desempenho desta aplicação, procedeu-se à ativação de um conjunto de índices, conforme indicado no Anexo B, aumentando assim a velocidade de determinadas pesquisas e, consequentemente, melhorando o desempenho.

A terceira etapa consistiu na instalação da aplicação ML. Após a instalação da aplicação UBA, constatou-se que o módulo de *Machine Learning* requeria uma instalação à parte. Na prática, este módulo é entendido como um a aplicação à parte, a qual pode ser instalada a partir da aplicação UBA. A sua disponibilização também é gratuita, tendo sido instalada conforme Anexo C. A versão desta última é coincidente com a da aplicação UBA.

A etapa final deste processo de configuração traduziu-se na importação dos dados dos utilizadores da Altice Portugal, residentes nos servidores da (AD) desta, como forma de enriquecimento da informação contextual e identificação unívoca. Conforme referido no capítulo de revisão do estado da arte, o protocol *Ligthweight Directory Access Protocol* (LDAP) pode ajudar a identificar a atividade de um dado utilizador ao longo de múltiplos e diferentes *logs*, pois este pode ser descrito nestes diferentes *logs* usando diferentes atributos. A aplicação UBA consegue agregar (*coalescing*) todas as atividades de utilizador, em todos os sistemas, num único utilizador. Como tal, esta etapa reveste-se de uma elevada importância, devendo a definição dos atributos da AD a considerar ser vista como essencial. Neste sentido, foram selecionados os seguintes atributos:

- Nome de Utilizador - atributo unívoco (pode funcionar como chave primária), na AD como *Logon Name*;
- Nome - atributo não unívoco que resulta da concatenação dos atributos *First Name* e *Last Name* da AD;
- Email - atributo unívoco na AD como *Email*;
- Departamento - atributo não unívoco na AD como *Department*;
- Cargo - atributo não unívoco na AD como *Job Title*.

Assim, os atributos que serão utilizados para executar o *User Coalescing* serão os atributos unívocos Nome de Utilizador e Email, enquanto que os restantes serão utilizados como atributos de visualização. O Anexo D detalha o processo de importação e configuração de utilizadores realizado, o qual segue a seleção acima. É importante dizer também que este processo foi realizado a partir da utilização de ficheiros CSV, os quais continham dados de utilizadores pertencentes a quatro dos domínios existentes na Altice Portugal. Apesar de a aplicação UBA permitir a interligação com servidores da AD, via protocolo LDAP, por dificuldades técnicas esta solução foi abandonada, e utilizada a solução alternativa via ficheiro CSV. A desvantagem desta última reside na incapacidade de implementar um mecanismo de atualização automática (terá de ser feito manualmente) dos dados desses utilizadores já carregados. Contudo, no âmbito deste projeto, não se considera relevante esta desvantagem ora apontada. Os dados depois de importados vão popular uma tabela de referência, a qual é lida pela aplicação UBA que, posteriormente, replica e guarda na base de dados interna. Para consultar os dados da tabela de referência pode ser realizado o procedimento descrito no Anexo E.

4.2 Eventos

Conforme explicado no Capítulo 3, Secção 3.2, os eventos de segurança recolhidos das estações de trabalho *Microsoft Windows* são armazenados nos diferentes coletores WEC, sendo depois filtrados, normalizados e reencaminhados pelo *NXLog* para o *QRadar*, com recurso ao protocolo *syslog*. Apesar de não fazer parte do âmbito deste projeto a configuração e subscrição destes eventos, este processo tem um impacto direto e significativo. Como tal, importa garantir que os eventos de segurança identificados como necessários a este projeto estão devidamente subscritos nestas diferentes etapas de recolha e encaminhamento. Este trabalho é desenvolvido e disponibilizado pela Altice Portugal. Assim, e na sequência dos casos de uso detalhados no Capítulo 3, Secção 3.3, foram subscritos os eventos indicados abaixo:

- Evento 4624 - Autenticação de uma conta com sucesso;
- Evento 4625 - Falha na autenticação de uma conta;
- Evento 4648 - Tentativa de autenticação usando credenciais explícitas;
- Evento 4778 - Restabelecimento de sessão numa estação *Windows*;
- Evento 4825 - Negação de acesso, a um dado utilizador, do ambiente de trabalho remoto.

Após garantida esta primeira pré-condição de subscrição e envio dos eventos para o *QRadar*, estes serão recebidos num formato *raw*. Para permitir a sua interpretação, análise e normalização, o *QRadar* recorre a um *Device Support Module* (DSM), habitualmente designado por *parser*. Este permite resolver alguns problemas de transformação, bem como adicionar outras transformações e campos de dados personalizáveis. Desta forma, torna-se também necessário validar que a informação dos eventos acima identificados está a ser corretamente extraída, e associada aos campos corretos. Por outro lado, e considerando que nem todos os campos de um evento no *QRadar* serão relevantes para este projeto, procurou-se identificar os campos, no Anexo F, que terão um potencial interesse e relevância neste projeto, possibilitando assim focalizar a validação dos dados dos eventos acima referidos.

Neste âmbito, foi monitorizada a receção de eventos, fazendo análises por amostragem, tendo-se constatado algumas incorreções, conforme indicado na tabela 4.1.

Tabela 4.1: Inconformidades detetadas nos eventos subscritos

ID Evento	Inconformidades Detetadas	Ações Tomadas
4624 e 4625	Campo Username não apresentava o nome de utilizador, apesar de este constar no evento, no formato *raw*	Corrigida a expressão regular que permite a extração deste dado
	Detetados eventos que mesmo no formato *raw* vêm muito incompletos (trazem apenas grupos data-hora do log, nome da máquina, domínio e ID Evento)	Estes eventos terão logicamente todos os outros campos a nulo, logo serão filtrados do conjunto de dados de análise
	Detetados eventos cujo nome de utilizador corresponde a um utilizador local (da máquina), nomeadamente Guest	Apesar de estes eventos estarem corretamente extraídos, não permitem uma atribuição, pelo que irão gerar falsos positivos. Estes serão filtrados do conjunto de dados de análise

4625	Detetados eventos em que o nome de utilizador era coincidente com o nome da estação de trabalho	Estes eventos deverão ser descartados do conjunto de dados de análise
	Campo Status Code não apresentava o código de estado, apesar de este constar no evento, no formato *raw*	Corrigida a expressão regular que permite a extração deste dado
	Campo Sub Status Code inexistente. Verificava-se a existência deste campo no evento no formato *raw*	Criado o campo respetivo e adicionada a expressão regular que permite a extração deste dado
4778	No período de monitorização não foram detetados eventos deste tipo. Após análise mais detalhada, constatou-se que os eventos deste tipo não são gerados nas estações de trabalho e, consequentemente, não são enviados para o *QRadar*	A utilização dos eventos com este ID foi descartada
4825	Campo Source Workstation não apresentava o nome da estação de trabalho, apesar de este constar no evento, no formato *raw*	Corrigida a expressão regular que permite a extração deste dado

Para finalizar esta fase de verificação e validação dos eventos, importa dizer que o *QRadar* possibilita a configuração de diferentes períodos de retenção dos dados de eventos. Por omissão, e após a instalação, esse período de retenção é configurado para um mês. Contudo, e tendo em conta os cenários a avaliar neste, os quais serão detalhados mais à frente, este mesmo parâmetro foi alterado, passando o período de retenção de eventos a ser de dois meses (sessenta dias).

4.3 Implementação

No Capítulo 3, Secção 3.3, foram definidos e apresentados três casos de uso de interesse para a organização Altice Portugal. No entanto, enquanto que o caso de uso *Tentativa de Ataque de Força Bruta* pode ser modelado deterministicamente, de acordo com o nível de apetite ao risco da organização, os outros dois casos de uso requerem modelos em que o comportamento dos utilizadores seja aprendido e, posteriormente, detetado qualquer desvio à atividade classificada como normal. Deste modo, foi considerado que o caso de uso *Tentativa de Ataque de Força Bruta* é mais vantajoso de modelar, numa relação de esforço-benefício, através da utilização de regras com aplicação de níveis de atividade ou tentativa de atividade de autenticação não desejada para a organização.

Como já descrito anteriormente, a aplicação ML disponibiliza, por omissão, um conjunto de modelos de casos de uso genéricos. Contudo, de acordo com os dois casos de uso a implementar, e considerando a sua especificidade, terão de ser criados novos modelos personalizados. Esta, aliás, é desde já uma funcionalidade desta aplicação. Para criar novos modelos basta aceder à página de configurações da secção *Machine Learning*, clicando depois no botão *Criar Modelo*. Após esta ação, é aberta uma janela com dois separadores, nos quais é apresentado um conjunto de campos, que é configurável pelos utilizadores. Neste sentido, de seguida descrevem-se os campos referidos:

- Separador *Model Definition* - este separador, como o nome indica, destina-se a definir o modelo, ou seja, a definir a *query* que será utilizada para popular o modelo de *Machine Learning*. A figura 4.1 ilustra os campos necessários a esta definição. No final, é apresentado um re-

sumo dos parâmetros configurados, sendo de salientar que todos os modelos apresentam uma granularidade temporal de uma hora, sem possibilidade de alteração deste parâmetro.

- Propriedade - corresponde à propriedade pertencente a um evento do QRadar, cujo valor será utilizado para construir o modelo. Só permite a definição de uma e uma só propriedade;
- Função - função que permite agregar e manipular os dados que são extraídos da base de dados *Ariel*, num determinado período de tempo (1h). São disponibilizados cinco tipos distintos de funções que podem ser utilizados: soma, média, máximo, mínimo e contador único;
- Filtro de Pesquisa *Ariel Query Language* (AQL) [17] - componente que permite a definição de um filtro para restringir o âmbito do modelo a um conjunto específico de dados.

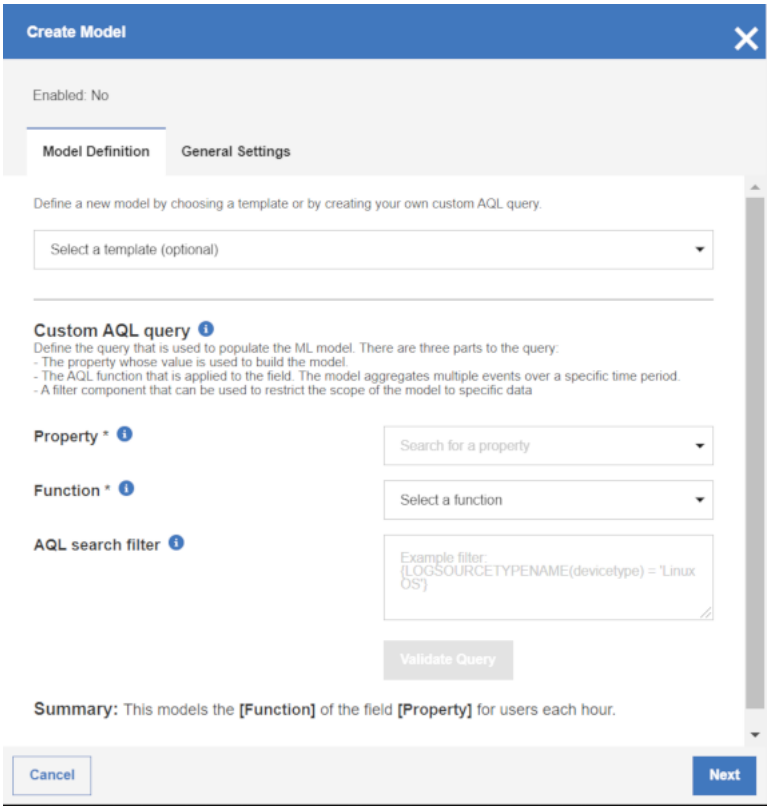

Figura 4.1: Criação de novo modelo Machine Learning Settings - Passo 1: Model Definition

- Separador *General Settings* - este separador permite definir um conjunto de parâmetros essenciais à elaboração do modelo, conforme se observa na figura 4.2, nomeadamente:
  - Nome - permite a atribuição de um nome ao modelo, o qual é utilizado para identificar este no gráfico, linha temporal e evento que é gerado quando uma anomalia é detetada;
  - Descrição - permite a introdução de uma descrição textual referente ao modelo criado;
  - Valor de Risco do *Sense Event* - corresponde à quantidade que será adicionada ao resultado de risco do utilizador quando um *Sense Event* é despoletado. Por omissão, é

[17] *Ariel Query Language* - linguagem estruturada, utilizada para comunicar com a base de dados *Ariel*, possibilitando, assim, pesquisar e manipular eventos e *flows*.

definido o valor de 5, sendo aceites valores entre 0 e 10000;

- Escalar Valor de Risco - campo binário (ativo ou não) que confere a opção de poder escalar o valor de risco de cada *Sense Event* num fator de 1 a 10, baseado no desvio em relação à atividade normal;
- Intervalo de Confiança para despoletar Anomalias - corresponde à definição do limite utilizado para despoletar um *Sense Event* quando o modelo avalia os dados de um utilizador. Este campo permite a introdução de valores de 0.5 a 0.999999;
- Período de Retenção de Dados - este campo é responsável pela definição do número de dias que os dados de um utilizador são retidos para o referido modelo. Podem ser definidos valores de 30 a 90 dias;
- Gráfico na Página de Detalhes do Utilizador - campo binário (ativo ou não) responsável por mostrar os resultados deste modelo sob a forma de gráfico, na página de perfil de um dado utilizador.

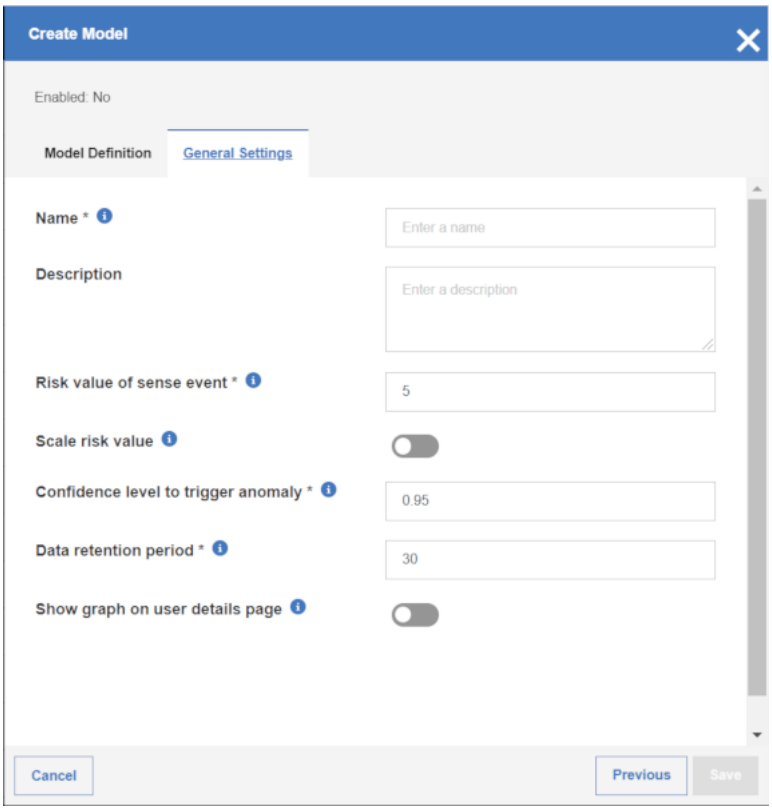

Figura 4.2: Criação de novo modelo Machine Learning Settings - Passo 2: General Settings

A implementação dos casos de uso identificados anteriormente assentou em quatro variantes de configuração, que diferiam entre si ao nível dos valores utilizados para configurar tanto os parâmetros do intervalo de confiança, como o período de retenção de dados, simultaneamente. A criação destas variantes tem como objetivo estudar quais os valores mais adequados a estes parâmetros numa implementação em produção, procurando minimizar o número de falsos positivos e maximizando os verdadeiros positivos. Para tal, consideram-se como gama de valores admissíveis para este caso de estudo, um valor de 30 ou 60 dias para o parâmetro período de retenção de dados, e um valor de 95% ou 99% para o parâmetro intervalo de confiança. Devido a esta necessidade de eventos por períodos de pelo menos 60 dias, foi necessário alterar a configuração do *QRadar* referente ao período de retenção de eventos. Por

omissão, este período é de um mês, tendo sido alterado para dois meses. Adicionalmente, decidiu-se adicionar um outro caso de estudo, o qual permitirá avaliar o desempenho relativo entre os casos de uso personalizados e um caso de uso genérico disponibilizado pela aplicação ML. Considerando o domínio dos eventos associados aos casos de uso definidos na Secção 3.3 a escolha de caso de uso genérico recaiu sobre o modelo *Authentication Activity*. Contudo, neste caso de uso, e existindo uma única versão do mesmo, não é possível estudar as variações dos parâmetros acima, como proposto para os caso de uso personalizados. Como tal, optou-se por estudar as configurações padrão sugeridas pela *IBM*. A tabela 4.2 apresenta uma relação dos vários casos de uso e suas variantes implementadas.

Tabela 4.2: Variantes de estudo dos diferentes casos de uso

Nome	ID Caso Uso	Intervalo Confiança	Período Retenção Dados
Autenticação Estação Trabalho Não Habitual	U1_01	95%	30 dias
	U1_02	95%	60 dias
	U1_03	99%	30 dias
	U1_04	99%	60 dias
Autenticação em Períodos Não Habituais	U2_01	95%	30 dias
	U2_02	95%	60 dias
	U2_03	99%	30 dias
	U2_04	99%	60 dias
Authentication Activity	U3	95%	30 dias

De seguida são descritas as implementações realizadas dos três casos de uso, considerando os parâmetros que são comuns a todas as variantes. É importante referir também que em todos os casos de uso os utilizadores ativos são monitorizados de forma contínua. Se um utilizador não apresentar qualquer registo de atividade durante 28 dias, o utilizador e os seus dados são removidos do modelo. Contudo, se o utilizador voltar a estar ativo novamente, então ele regressará com um novo utilizador.

Caso Uso 1 - Autenticação em Estação de Trabalho Não Habitual

Neste primeiro caso de uso procura-se detetar tentativas de autenticação, bem sucedidas ou não, por um utilizador numa estação de trabalho não habitualmente utilizada por este. A figura 4.3 ilustra a implementação do referido modelo.

Como se observa, foi definido como propriedade o atributo *Source Workstation*, o qual corresponde à estação de trabalho onde decorre a tentativa de autenticação. Como função, foi escolhida a função de contagem única/distinta da propriedade anterior. Posteriormente, definiu-se o filtro de pesquisa AQL, conforme consta no Anexo G, por forma a definir os dados que serão analisados pelo modelo. De referir que este filtro é igual para todos os casos de uso.

Na figura 4.4, é apresentada a definição das configurações gerais do modelo de caso de uso. Para além da definição do nome (neste caso corresponde ao ID do caso de uso para mais fácil identificação) e descrição, há a salientar a escolha do valor de risco do *Sense Event* de 20 pontos, não tendo sido ativada a opção de escalar esse mesmo valor. Por fim, foi ativada a opção que permite a visualização de um gráfico comportamental na página de detalhes do utilizador. Uma vez que os campos intervalo de confiança e período de retenção de dados serão alvo de estudo, conforme mencionado anteriormente, esta demonstração de implementação é feita apenas para o primeiro caso, a título de exemplo.

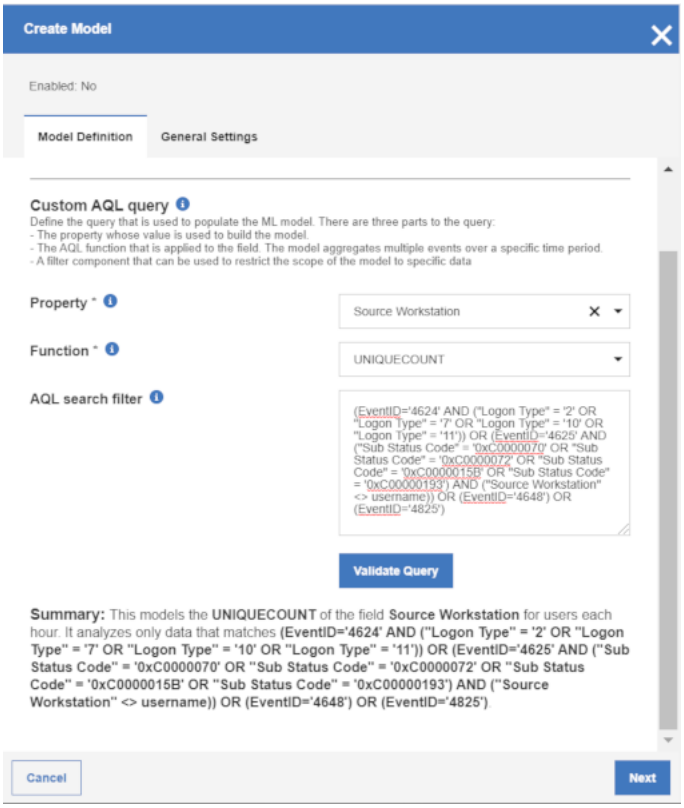

Figura 4.3: Caso Uso 1 - Passo 1: Definição do Modelo

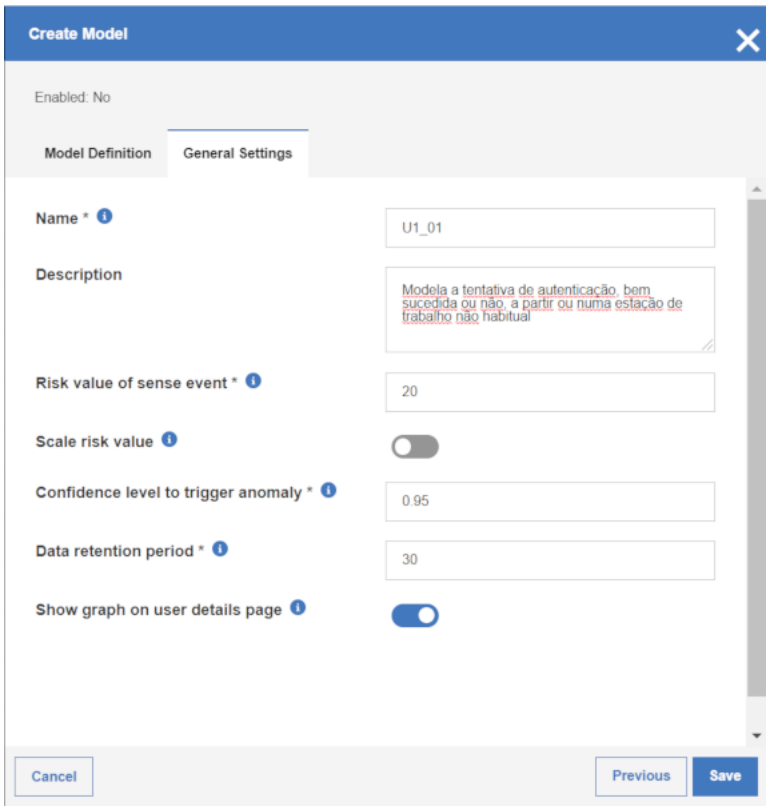

Figura 4.4: Caso Uso 1 - Passo 2: Definições Gerais

Caso Uso 2 - Autenticação em Períodos Não Habituais

Neste segundo caso de uso, o objetivo passa por detetar tentativas de autenticação, bem sucedidas ou não, por um utilizador num período que difere do habitualmente utilizado por este. Para este caso, como é possível observar na figura 4.5, foi utilizado o atributo *Event Hour*, o qual corresponde ao campo hora, não entrando em linha de conta com os minutos e segundos. Esta opção de apenas incluir o campo hora justifica-se pelo facto de que, se a granularidade fosse maior, incluindo minutos e segundos, poderia traduzir-se num aumento de falsos positivos não desejados. Para obter este atributo foi necessário criar um tipo de atributo personalizado, aplicando uma expressão regular para a selecção da hora, a partir do atributo *Log Source Time*. À semelhança do caso de uso anterior, foi escolhida a função de contagem única/distinta da propriedade *Event Hour*. O filtro de pesquisa aplicado foi, como já referido e explicado, igual ao caso de uso anterior.

Contudo, após realização de alguns testes iniciais concluíu-se que este atributo e função associada não apresentavam o comportamento desejado. Mais, esta propriedade era do tipo *String*, pelo que não permitia ser utilizada por funções numéricas (máximo, mínimo, média ou soma). Assim, optou-se por substituir a propriedade anterior pelo atributo *Event Count*[18] e, a respetiva função pela *Soma*, ou seja, num período não habitual a soma total de eventos deverá corresponder a zero. Foi também considerada a hipótese de utilizar da função *Máximo*, mas considerando a definição deste atributo, a opção mais correta seria fazer uso da função *Soma*. O filtro de pesquisa foi mantido sem qualquer alteração.

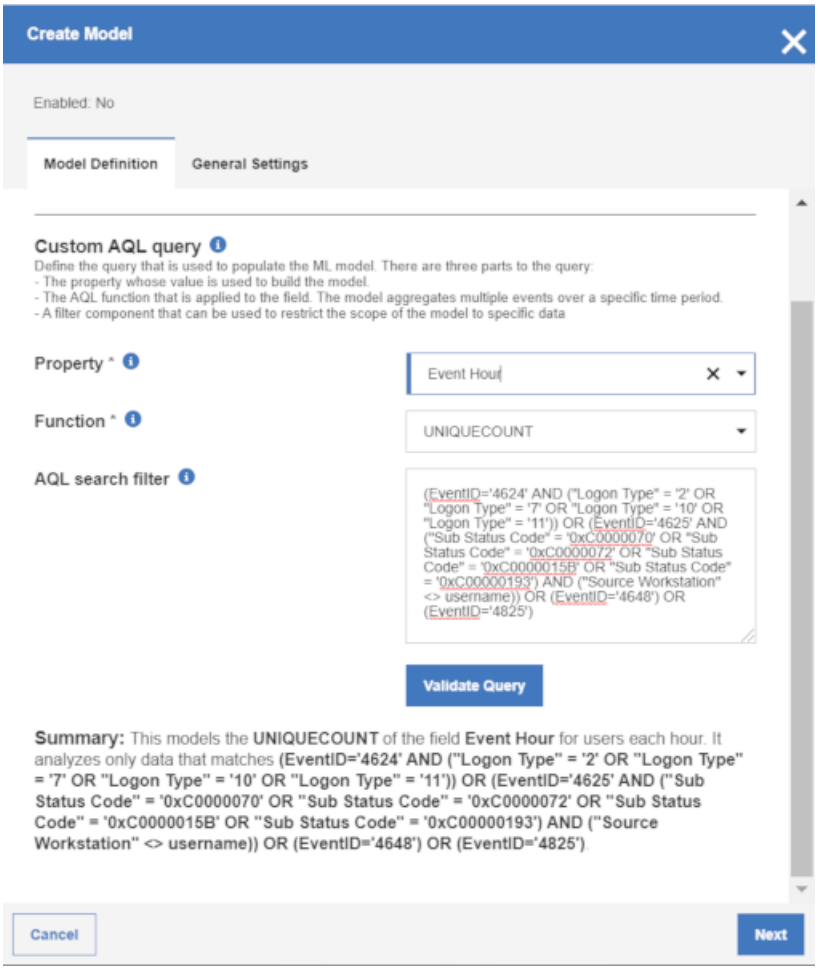

Figura 4.5: Caso Uso 2 - Passo 1: Definição Modelo

[18]Especifica o número total de eventos que são agrupados num dado evento normalizado. Os eventos são agrupados quando possuem o mesmo tipo e origem, com destino ao mesmo endereço IP e são detetados num curto espaço de tempo.

Na definição das configurações gerais, visíveis na figura 4.6, importa apenas realçar o facto de ter sido escolhido um valor de 30 pontos para o valor de risco do *Sense Event*. Os restantes parâmetros são semelhantes aos do caso de uso anterior e, por isso, não serão novamente descritos.

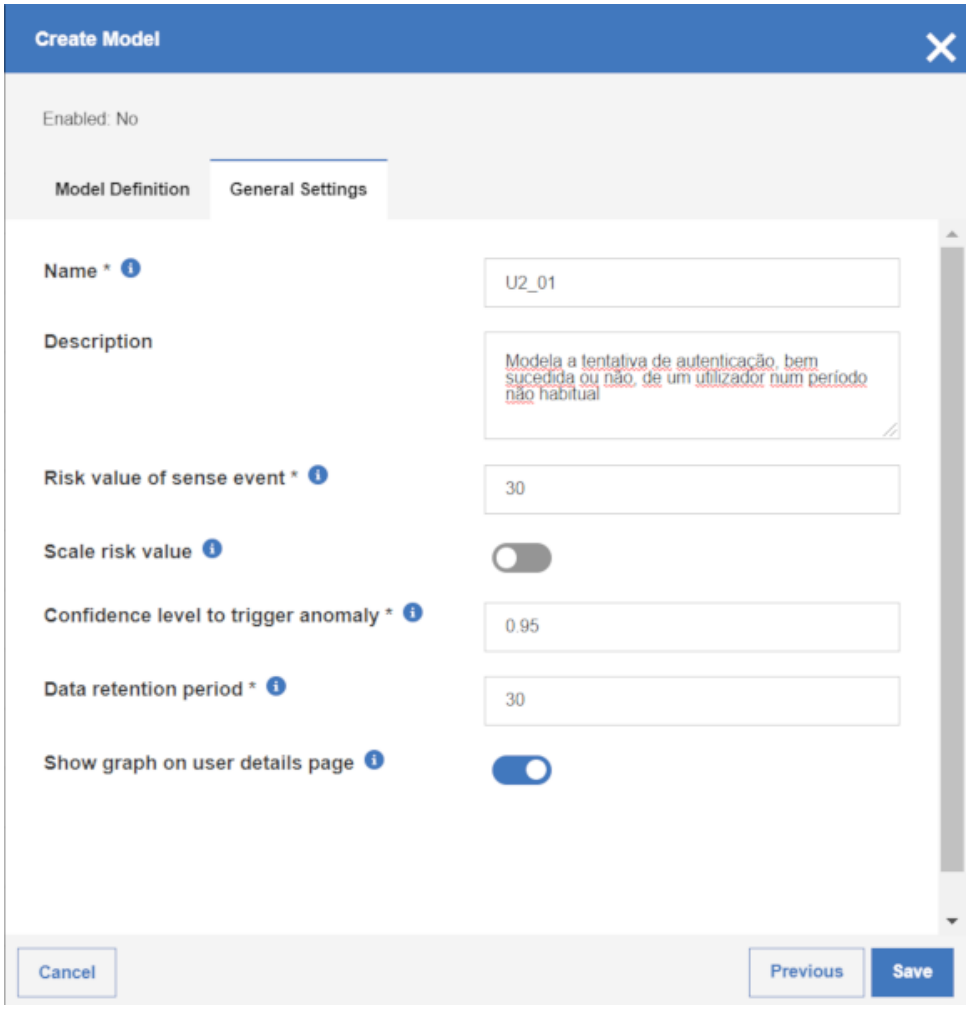

Figura 4.6: Caso Uso 2 - Passo 2: Definições Gerais

Caso Uso 3 - Authentication Activity

Este caso de uso enquadra-se num conjunto de casos de uso genéricos, desenvolvidos pela *IBM* e disponibilizados de forma a serem explorados e ativados, requerendo poucas ou nenhumas configurações. Este caso de uso em particular destina-se a fazer o seguimento da atividade dos utilizadores, considerando, para tal, todos os eventos que sejam categorizados com a categoria de alto nível de autenticação. Deste modo, ele consegue, assim, construir e criar um modelo de aprendizagem comportamental para cada hora do dia, à semelhança do que sucede com os casos de uso anteriores.

A figura 4.7 permite demonstrar as configurações que são possíveis efetuar nestes modelos pré-definidos. Conforme já referido, por questões de análise foram mantidos os valores padrão, com exceção do valor de risco do *Sense Event*, o qual foi alterado para 20 pontos.

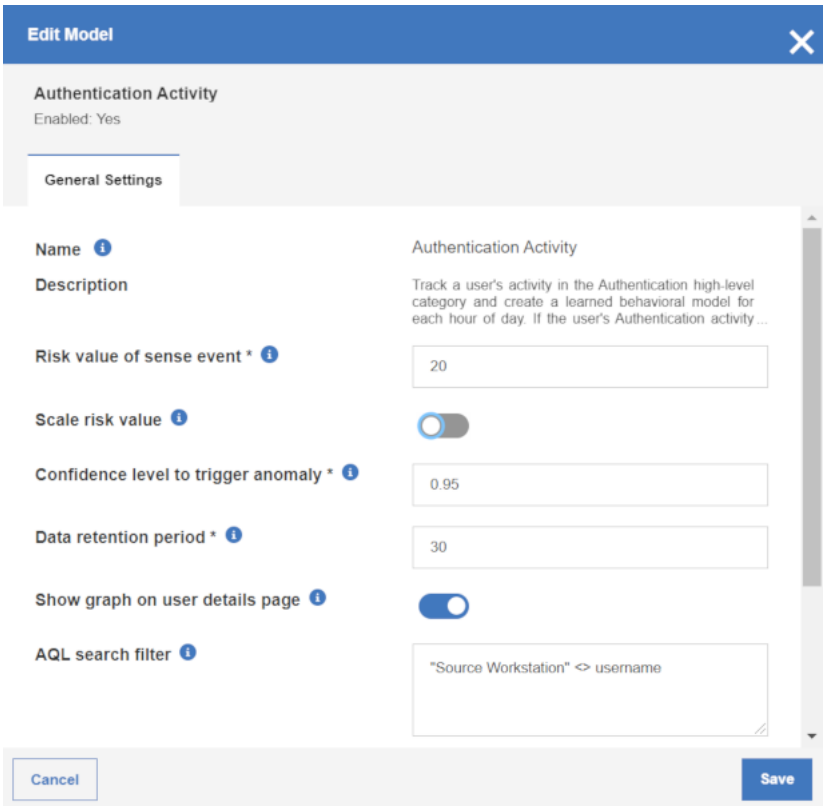

Figura 4.7: Caso Uso 3 - Authentication Activity

4.4 Testes

Esta secção tem como objetivo descrever a forma e os tipos de teste que foram realizados, de modo a permitir uma avaliação o mais completa e abrangente possível de todas as diferentes variantes e casos de uso em análise. Os testes tiveram como população alvo os utilizadores da Altice Portugal, mais especificamente os que integram a Direção de Cyber Security e Privacidade (DCY).

De seguida, são descritos os quatro tipos diferentes de testes realizados, as suas variantes de execução, bem como os seus objetivos finais. Uma vez que os eventos de servidores não estão, presentemente, a ser recebidos, estes testes ficam limitados ao universo de estações de trabalho. Importa salientar ainda que as duas variantes de execução de cada teste devem ser realizadas fora do mesmo intervalo horário, para que os resultados de ambas não sejam afetados. A título de exemplo, se a primeira variante for realizada entre o período das 10h às 11h, a segunda variante só poderá ocorrer a partir das 11h.

Teste I

* Tipo de Teste: Autenticação em estação de trabalho atribuída, no período habitual de trabalho;
* Objetivo: este teste procura validar o aparecimento, ou não, de falsos positivos, isto é, se são gerados, ou não, alertas a partir de eventos que representam o comportamento habitual de um dado utilizador;
* Variantes de Execução: a execução das duas variantes não pode ocorrer dentro do mesmo intervalo horário;
    # *Variante I* - realizar uma autenticação bem sucedida na estação de trabalho atribuída, durante o seu período habitual de trabalho;

- # *Variante II* - realizar vinte ou mais autenticações bem sucedidas na estação de trabalho atribuída, no intervalo de uma hora, dentro do seu período habitual de trabalho.

Teste II

- ∗ Tipo de Teste: Autenticação em estação de trabalho não atribuída, no período habitual de trabalho;
- ∗ Objetivo: este teste destina-se a validar o correto funcionamento do caso de uso 1;
- ∗ Variantes de Execução: a execução das duas variantes não pode ocorrer dentro do mesmo intervalo horário;
  - # *Variante I* - realizar uma autenticação bem sucedida numa estação de trabalho que não lhe esteja atribuída, durante o seu período habitual de trabalho;
  - # *Variante II* - realizar cinco ou mais autenticações bem sucedidas e/ou cinco ou mais tentativas de autenticação mal sucedidas (ex: falhar a palavra-passe), sempre numa estação de trabalho distinta da que lhe esteja atribuída, no intervalo de uma hora, durante o seu período habitual de trabalho.

Teste III

- ∗ Tipo de Teste: Autenticação em estação de trabalho atribuída, fora do período habitual de trabalho;
- ∗ Objetivo: este teste destina-se a validar o correto funcionamento do caso de uso 2;
- ∗ Variantes de Execução: a execução das duas variantes não pode ocorrer dentro do mesmo intervalo horário;
  - # *Variante I* - realizar uma autenticação bem sucedida na estação de trabalho atribuída, fora do seu período habitual de trabalho;
  - # *Variante II* - realizar dez ou mais autenticações bem sucedidas e/ou dez ou mais tentativas de autenticação mal sucedidas (ex: falhar a palavra-passe), na estação de trabalho atribuída, no intervalo de uma hora, fora do seu período habitual de trabalho.

Teste IV

- ∗ Tipo de Teste: Autenticação em estação de trabalho não atribuída, fora do período habitual de trabalho;
- ∗ Objetivo: este teste destina-se a validar o correto funcionamento, em simultâneo, dos casos de uso 1 e 2;
- ∗ Variantes de Execução: a execução das duas variantes não pode ocorrer dentro do mesmo intervalo horário;
  - # *Variante I* - realizar uma autenticação bem sucedida numa estação de trabalho que não lhe esteja atribuída, fora do seu período habitual de trabalho;
  - # *Variante II* - realizar cinco ou mais autenticações bem sucedidas e/ou cinco ou mais tentativas de autenticação mal sucedidas (ex: falhar a palavra-passe), sempre numa estação de trabalho distinta da que lhe esteja atribuída, no intervalo de uma hora, fora do seu período habitual de trabalho.

Os procedimentos utilizados para os diferentes testes de utilizadores, bem como todos os modelos de questionários, podem ser encontrados no Anexo H. Os procedimentos adotados foram os mesmos para

todos os utilizadores: inicialmente foi enviado por correio eletrónico o guião para a execução dos testes, e, posteriormente à conclusão do mesmo, foi solicitado o preenchimento de um questionário referente a esse mesmo teste. Considerando que não é possível a execução dos testes de forma seguida, esta traduziu-se na forma mais adequada de registar os comentários dos utilizadores, visando assim permitir a comparação com os resultados obtidos através da análise do registo de atividade na aplicação *QRadar*.

Mormente, e conhecendo à partida a amostra de teste, foi explorada uma das funcionalidades do *QRadar*, nomeadamente a criação de uma *watchlist*. As *watchlists* traduzem-se num mecanismo que permite agregar um conjunto de utilizadores, por forma a especificar, para esses mesmos utilizadores, um conjunto de parâmetros, como são o caso do fator de escalonamento do risco e a prioridade de seguimento com *Machine Learning*. Este último, após análise, revela-se como um parâmetro fundamental, uma vez que ele determina a prioridade com que os utilizadores são seguidos pelo modelos analíticos de *Machine Learning*, ou seja, conforme referido pelo manual de utilizador da aplicação UBA [42], estes apresentam um limite máximo de utilizadores que podem ser seguidos (este valor não é conhecido nem divulgado publicamente), pelo que se for atribuída uma prioridade mais elevada aos utilizadores de uma dada *watchlist*, então eles serão sempre seguidos até que o limite máximo seja atingido. Neste contexto, foi então criada uma *watchlist*, de nome *DCY Watchlist*, para mais fácil monitorização e configuração dos utilizadores envolvidos nos testes, conforme procedimentos descritos no Anexo I. Esta foi definida com uma prioridade elevada pela razão explicada anteriormente. Simultaneamente, e como facilmente se depreende, esta configuração não garante que um dado utilizador seja seguido pelos modelos analíticos. Assim, e uma vez que o *QRadar* também permite forçar este seguimento, de forma manual, na página de detalhes de cada utilizador, optou-se também por realizar esta ação para todos os utilizadores presentes na *watchlist* anterior, conforme se detalha também no Anexo I.

Por outro lado, nesta fase dos testes recorreu-se também à utilização de *scripts* que permitiram automatizar o processo de injeção de eventos de autenticação no *QRadar*, conferindo uma maior consistência com os padrões de comportamento normal de trabalho, uma vez que estes foram necessariamente afetados devido à pandemia causada pelo vírus *SARSCoV-2*. Mais, estes *scripts* conferem também a possibilidade de executar as diferentes variantes dos testes definidos de uma forma fácil e expedita, permitindo, assim, fazer um acompanhamento gradual dos modelos implementados, testando as condições limite. Neste sentido, foram desenvolvidos três ficheiros distintos, programados em *bash*, os quais podem ser encontrados no Anexo J. Enquanto que os dois primeiros *scripts* visam contribuir para o processo de injeção de eventos de autenticação, tendo a sua execução sido configurada com recurso ao escalonador de tarefas do sistema operativo, o terceiro *script* destina-se a permitir a geração de eventos anómalos, qualitativa e quantitativamente. Deste modo, de seguida, detalha-se as funções e execução de cada um dos *scritps*:

- *Script* para Geração de Eventos do Tipo 4624 (autenticação de uma conta com sucesso) - este *script* permite gerar um número aleatório de eventos de autenticação do tipo 4624, compreendido entre 1 e 10, por hora. Similarmente, sempre que é executado, é escolhido um número aleatório entre 1 e 55, como o minuto de ocorrência dos eventos. Este *script* é executado, de forma automática, uma vez por hora nos dias úteis (segunda a sexta-feira), no período das 8h às 19h, apenas para o utilizador *xfcta16*, procurando assim padronizar o horário de trabalho deste utilizador;
- *Script* para Geração de Eventos do Tipo 4625 (falha de autenticação de uma conta) - este *script* permite gerar um número aleatório de eventos de autenticação do tipo 4625, compreendido entre 1 e 3, sendo o envio do número definido anteriormente seguido sempre por um evento do tipo 4624 (pretende simular a tentativa de autenticação, com um máximo de 3 falhas, seguido de uma autenticação bem sucedida). À semelhança do *script* anterior, é escolhida a hora e minuto de envio dos eventos anteriores descritos. Este *script* é executado, de forma automática,

uma vez por dia, nos dias úteis (segunda a sexta-feira), a partir das 8h, também apenas para o utilizador *xfcta16*;

- *Script* para Geração de Eventos Anómalos - *script* que é executado de forma manual, permitindo ao utilizador que o corre definir o tipo de eventos que pretende gerar (4624 ou 4625), bem como o número de eventos a enviar. Dependendo do tipo de teste para o qual está a ser utilizado, poderá ser necessário ser alterado tanto o utilizador como a estação de trabalho de origem.

Capítulo 5

Resultados e Discussão

5.1 Resultados

Neste capítulo, após definição e implementação dos modelos associados a cada um dos casos de uso no capítulo anterior, será efetuada a apresentação e análise dos resultados obtidos, com base no cumprimento dos objetivos propostos de cada caso de uso e, consequentemente, nos resultados obtidos.

Os resultados apresentados neste capítulo foram obtidos através da condução dos vários testes, e respetivas variantes, identificados no Capítulo 4, Secção 4.4, fazendo uso também dos guiões e questionários identificados no Anexo H. Os resultados obtidos respeitaram um período de aprendizagem dos modelos superior a 60 dias.

Previamente à descrição e discussão dos resultados, importa referir uma desvantagem da utilização de modelos personalizados, em contraste com a utilização dos modelos disponibilizados por omissão: sempre que um dado modelo personalizado, ou não, deteta uma anomalia comportamental, é despoletado um evento, o qual pode ser visualizado no separador *Log Activity*. Enquanto que, no caso dos modelos por omissão, a identificação do modelo em questão é executada de forma clara e perceptível para o analista, no caso dos modelos personalizados, essa identificação é pouco explícita e clara, isto é, apesar de durante a criação destes modelos ser atribuído um nome ao mesmo, este não é utilizado para proceder a essa identificação. Para tal, é utilizado um campo *usecaseUUID* gerado no momento da criação do modelo, não conhecido pelo utilizador, o qual fica armazenado no ficheiro de configuração deste. Assim, para garantir a correção dos resultados apresentados seguidamente, foram extraídos os identificadores de cada modelo, possibilitando assim elaborar sempre um mapeamento entre os modelos que despoletam uma dada anomalia e os eventos gerados.

5.1.1 Teste I - Autenticação em Estação de Trabalho Atribuída no Período Habitual de Trabalho

O primeiro teste visa a realização de autenticações bem sucedidas, na estação de trabalho atribuída a um dado utilizador, durante o seu período habitual de trabalho. O objetivo deste teste passa por validar o correto funcionamento dos modelos personalizados implementados, considerando, para tal, que não são gerados falsos positivos. Este teste foi dividido em duas variantes, sendo que na primeira variante os utilizadores executaram apenas uma autenticação bem sucedida, enquanto que na segunda executaram vinte ou mais autenticações bem sucedidas, no intervalo de uma hora.

A primeira variante foi realizada por catorze utilizadores, sendo que apenas para treze destes foi possível confirmar a correta execução do teste, uma vez que a estação de trabalho de um utilizador se apresentava no estado de erro, não enviando, desse modo, quaisquer eventos para o *QRadar*. Para os utilizadores validados verificou-se que, todos sem exceção, não despoletaram qualquer anomalia para

todos os diferentes casos de uso em estudo, não sendo assim gerado qualquer falso positivo nesta variante. Desta forma, considera-se que esta variante foi executada com sucesso.

A segunda variante deste teste foi executada por dez utilizadores. Os resultados obtidos são mostrados na figura 5.1. O caso de uso U1 permite caracterizar o comportamento de um utilizador quanto ao número de estações de trabalho distintas que utiliza, sendo sempre independente do número de autenticações que produz nessas mesmas estações de trabalho. De forma análoga, o caso de uso U2 possibilita a caracterização do comportamento de um utilizador quanto às horas distintas em que este se autentica. Conforme se observa, tanto para todas as variantes do caso de uso U1, como para todas as variantes do caso de uso U2, não ocorreu qualquer deteção de anomalia associada ao comportamento dos utilizadores (teste realizado), resultado este que está de acordo com o esperado e que prova que os modelos estão corretos, não gerando falsos positivos. Já o caso de uso U3 é responsável por medir a atividade de autenticação. No presente teste, todos os utilizadores realizaram mais de vinte autenticações bem sucedidas no intervalo de uma hora, o que não traduz um comportamento habitual destes. Na figura 5.1 observa-se que seis dos dez utilizadores despoletaram uma anomalia para este caso de uso, com uma confiança de 100%, enquanto que os restantes quatro não despoletaram anomalia, pois a confiança era abaixo do definido (95%).

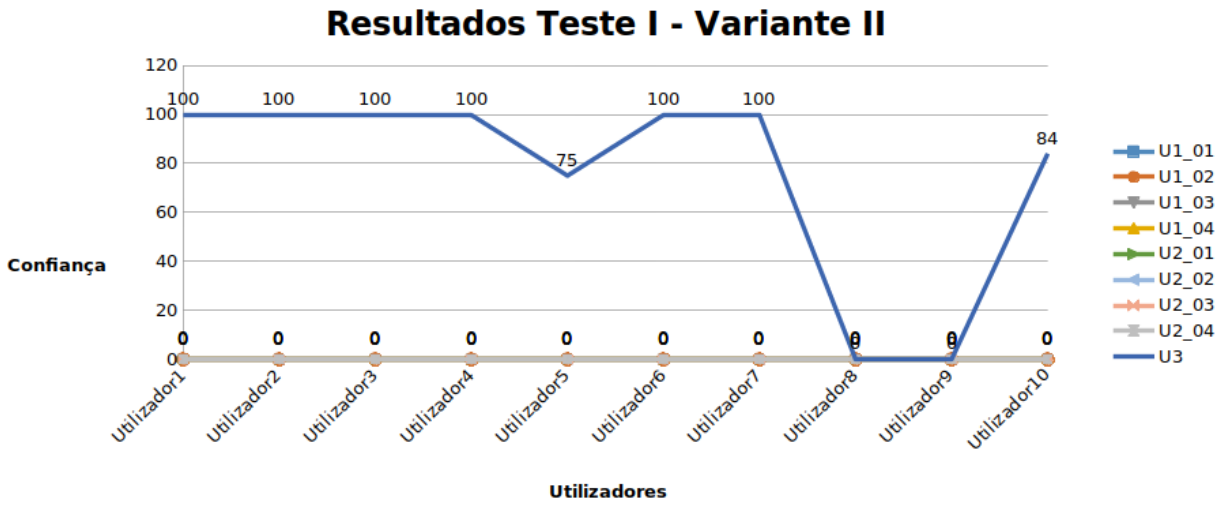

Figura 5.1: Resultados Teste I - Variante II

5.1.2 Teste II - Autenticação em Estação de Trabalho Não Atribuída no Período Habitual de Trabalho

O segundo teste visa a realização de autenticações, bem sucedidas ou não, executadas sempre em estações de trabalho distintas e diferentes da atribuída a um dado utilizador, durante o seu período habitual de trabalho. O objetivo deste passa por validar o correto funcionamento do caso de uso U1, bem como analisar quais os parâmetros que mais se adequam a este. Este teste, tal como o anterior, foi dividido também em duas variantes, sendo que na primeira variante os utilizadores executaram apenas uma autenticação bem sucedida, enquanto que na segunda executaram cinco ou mais autenticações bem sucedidas ou não, no intervalo de uma hora.

A execução da primeira variante deste teste foi realizada por doze utilizadores, tendo sido obtidos os resultados ilustrados nas figuras 5.2 e 5.3. A primeira figura permite estabelecer um contraste entre o comportamento aprendido e o comportamento atual para cada caso de uso, enquanto que a segunda mostra o nível de confiança atribuído à situação testada.

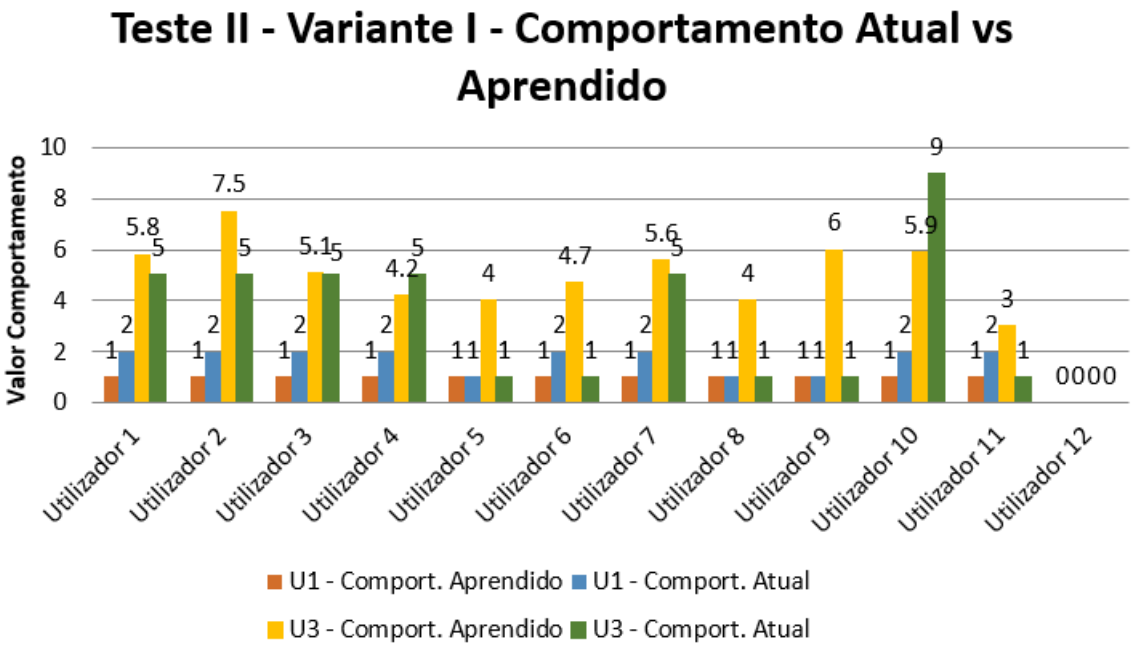

Figura 5.2: Teste II - Variante I - Comportamento Atual vs Aprendido

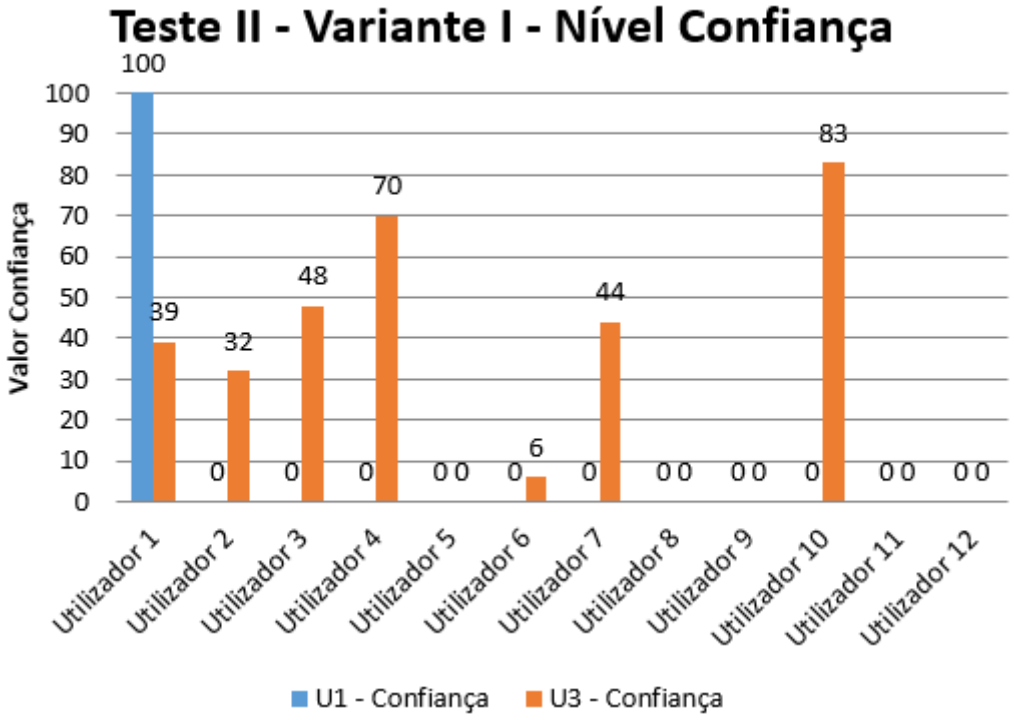

Figura 5.3: Teste II - Variante I - Nível de Confiança

Para o caso de uso U1, foi registado um comportamento igual para todas as variantes, optando-se, assim, por agrupar os resultados destas apenas numa única série, representada pelo caso de uso U1. Considerando que o modelo deste caso de uso elabora uma contagem distinta de estações de trabalho utilizadas, e que um utilizador só trabalha numa dada estação, o simples facto de um utilizador se autenticar numa estação de trabalho que não é a sua habitual deveria ser suficiente para este modelo despoletar uma anomalia. Deste modo, podem ser assim resumidos os resultados:

- Apenas um utilizador (Utilizador 1) registou como anomalia comportamental o presente teste, apresentado uma confiança de 100% para todas as variantes;
- Sete utilizadores (Utilizador 2, 3, 4, 6, 7, 10, 11) demonstram na figura 5.2 uma alteração significativa ao seu padrão comportamental, apresentando para todas as variantes uma confiança de 0%, pelo que não existe registo de qualquer anomalia;
- Três utilizadores (Utilizador 5, 8, 9) não possuem qualquer alteração seu ao padrão comportamental, não sendo por isso gerada qualquer anomalia e apresentando igualmente uma confiança de 0%;
- Para um dos utilizadores (Utilizador 12), não foi possível sequer detetar o evento de autenticação produzido por este. Após análise detalhada, confirmou-se que o mesmo fez uso de uma estação de trabalho pessoal, a qual, apesar de se ligar à rede corporativa da Altice Portugal, não obedece à política de recolha de eventos implementada.

O caso de uso U3 é responsável pela medição da atividade de autenticação. Considerando que este teste é realizado no período habitual de trabalho, é expectável os utilizadores apresentarem registo de diferentes tentativas de autenticação, pelo que não será suposto este modelo ser responsável pelo despoletar de qualquer anomalia, caso o comportamento não varie muito do habitual. De acordo com os resultados obtidos, verifica-se que o comportamento atual, na maioria dos utilizadores, não é substancialmente diferente do comportamento aprendido, pelo que não foi gerada qualquer anomalia.

A segunda variante deste teste procura, de forma semelhante à variante anterior, explorar a capacidade de deteção de anomalias por utilização de estações de trabalho não habituais, só que utilizando um número maior de estações de trabalho diferentes para realizar as tentativas de autenticação, visando, assim, aferir melhor a capacidade de deteção do modelo. Esta variante foi executada por dez utilizadores, sendo que, para o caso de uso U1, se obteve os resultados presentes nas figuras 5.4 e 5.5 e descritos de seguida:

- Apenas um utilizador (Utilizador 1) registou como anomalia comportamental o presente teste, apresentado uma confiança de 100% para todas as variantes;
- Os restantes utilizadores apresentam na figura 5.4 uma alteração significativa ao seu padrão comportamental, mas possuem para todas as variantes uma confiança de 0%, pelo que não existe registo de qualquer anomalia.

Quanto ao caso de uso U3, tal como na variante anterior, não foram detetadas diferenças significativas de comportamento dos utilizadores, pelo que, conforme esperado, este modelo não deu origem a qualquer anomalia.

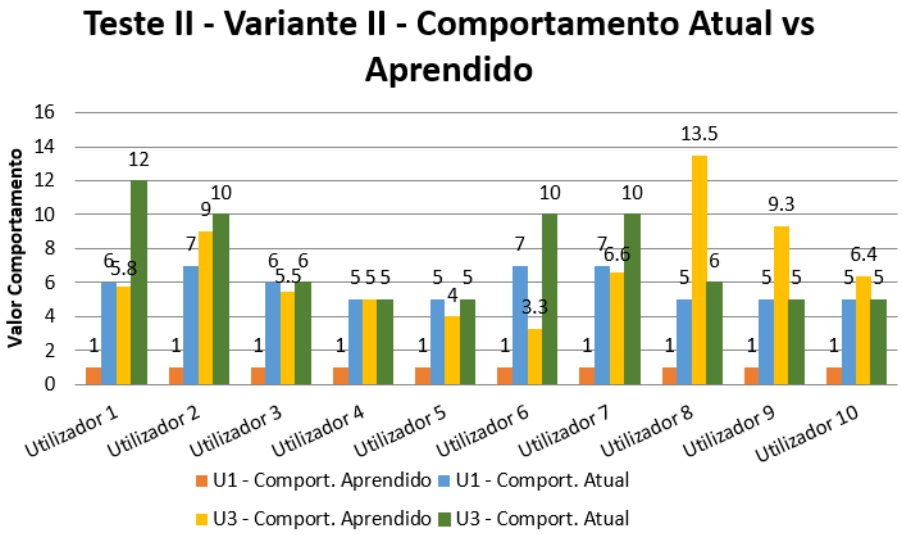

Figura 5.4: Teste II - Variante II - Comportamento Atual vs Aprendido

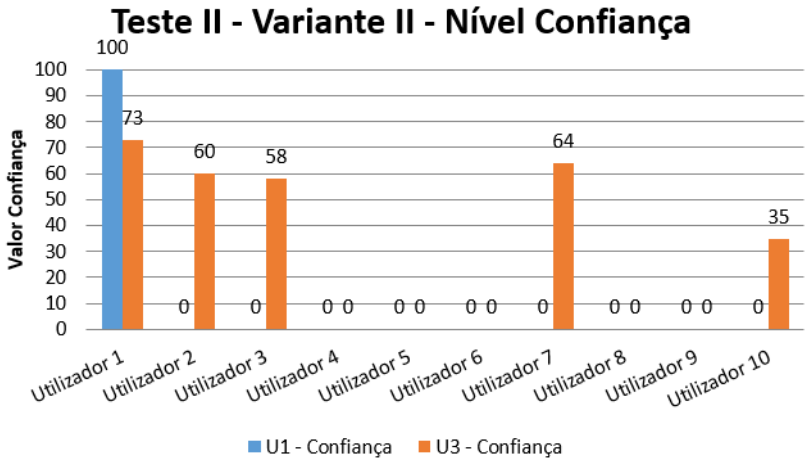

Figura 5.5: Teste II - Variante II - Nível de Confiança

5.1.3 Teste III - Autenticação em Estação de Trabalho Atribuída Fora do Período Habitual de Trabalho

O terceiro teste envolveu a realização de autenticações, bem sucedidas, na estação de trabalho atribuída a um dado utilizador, fora do seu período habitual de trabalho. O objetivo deste teste passava por validar o correto funcionamento do caso de uso U2, bem como analisar quais os parâmetros que mais se adequam a este. Este teste, de igual modo aos anteriores, foi dividido em duas variantes, sendo que na primeira variante os utilizadores executaram apenas uma autenticação bem sucedida, enquanto que na segunda executaram dez ou mais autenticações bem sucedidas ou não, no intervalo de uma hora.

A primeira variante deste teste foi realizada por dez utilizadores, tendo sido obtidos os resultados que se observam nas figuras 5.6 e 5.7.

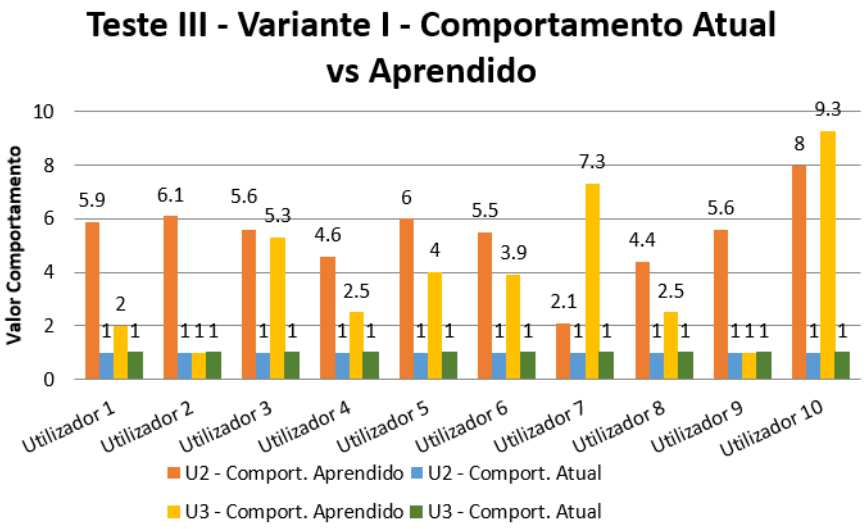

Figura 5.6: Teste III - Variante I - Comportamento Atual vs Aprendido

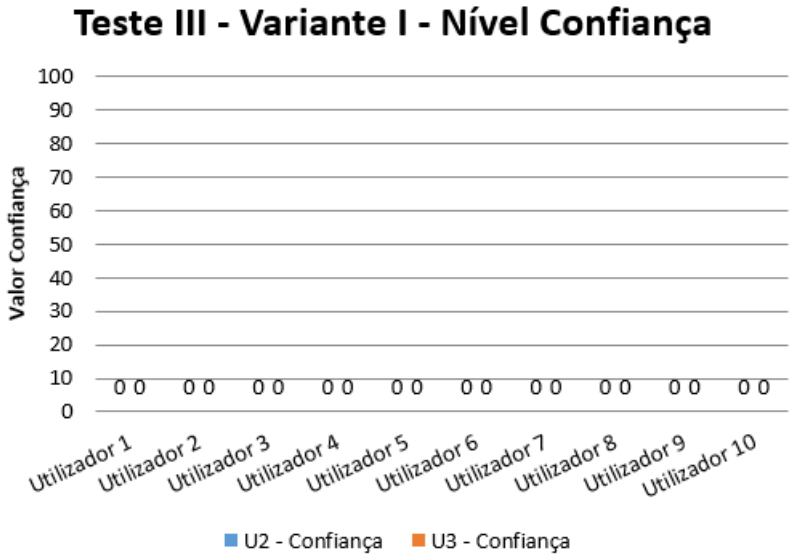

Figura 5.7: Teste III - Variante I - Nível de Confiança

Para o caso de uso U2, foi também registado um comportamento igual para todas as variantes, optando-se assim por agrupar novamente os resultados destas apenas numa única série, representada pelo caso de uso U2. O modelo deste caso de uso elabora uma soma dos eventos de autenticação que ocorrem numa dada hora, para cada utilizador. Neste sentido, se um dado utilizar realizar autenticações num período não habitual, onde o comportamento esperado seja zero, então deverá ser originada uma anomalia. Nesta primeira variante, observa-se que o comportamento atual dos utilizadores, tanto para o caso de uso U2, como para o caso de uso U3, é sempre igual (corresponde a uma autenticação, tal como programado no teste). No entanto, o comportamento aprendido por ambos os modelos é sempre igual ou superior ao comportamento atual, pelo que nunca é despoletada qualquer anomalia.

De igual modo, na segunda variante deste teste, observa-se que o comportamento atual dos utilizadores é sempre igual, tanto para o caso de uso U2, como para o caso de uso U3 (corresponde ao número de autenticações realizadas pelos utilizadores durante o teste). Apesar de se observar um contraste entre o comportamento atual desta variante e o da variante anterior, sendo este sempre superior ao comporta-

mento aprendido de ambos os casos de uso, verifica-se contudo que não foi originada qualquer anomalia comportamental pelos utilizadores testados. Estes resultados podem ser observados nas figuras 5.8 e 5.9.

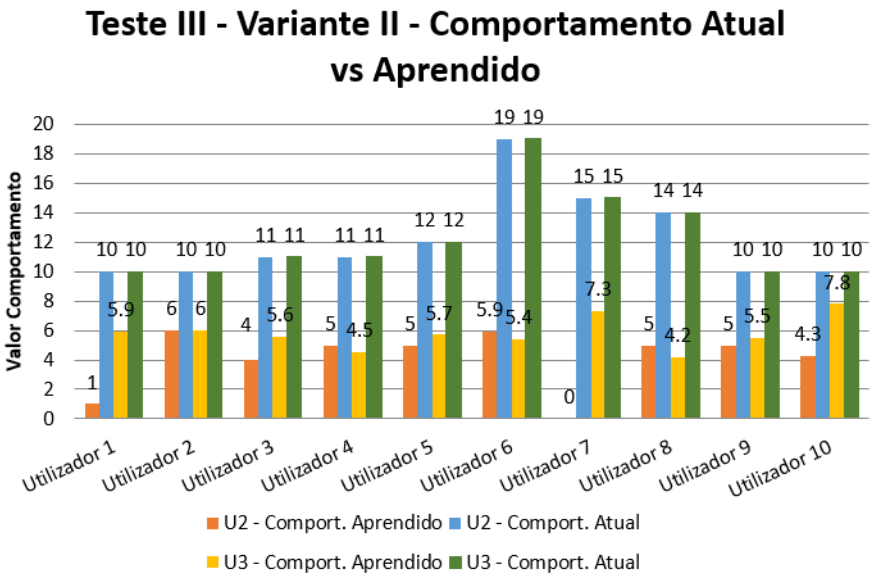

Figura 5.8: Teste III - Variante II - Comportamento Atual vs Aprendido

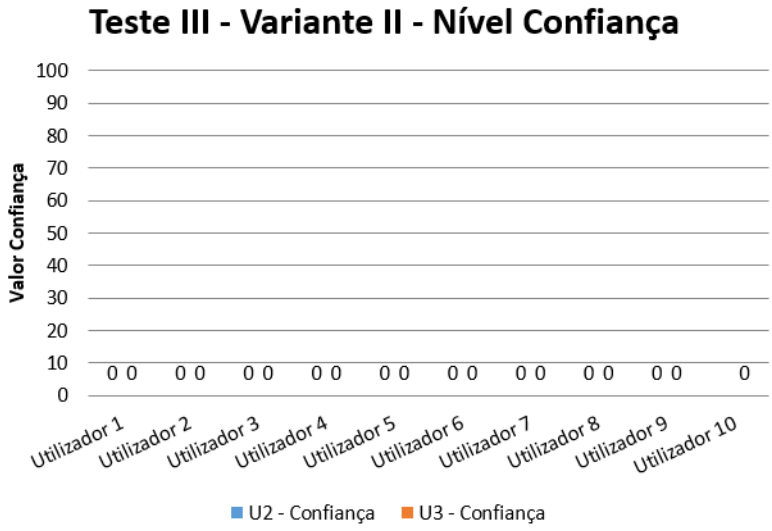

Figura 5.9: Teste III - Variante II - Nível de Confiança

5.1.4 Teste IV - Autenticação em Estação de Trabalho Não Atribuída Fora do Período Habitual de Trabalho

O quarto e último teste compreendeu a realização de autenticações bem sucedidas ou não, executadas sempre em estações de trabalho distintas e diferentes da atribuída a um dado utilizador, fora do seu período habitual de trabalho. O objetivo deste passava por validar o correto funcionamento, em simultâneo, dos casos de uso U1 e U2, bem como analisar quais os parâmetros que melhor se adequavam a estes. À semelhança dos testes anteriores, foi também dividido em duas variantes, sendo que na primeira variante os utilizadores executaram apenas uma autenticação bem sucedida, enquanto que na segunda executaram cinco ou mais autenticações, bem sucedidas ou não, no intervalo de uma hora.

Da execução da primeira variante deste teste obteve-se os resultados constantes nas figuras 5.10 e 5.11. Como se observa, para qualquer um dos três casos de uso testados, simultaneamente, nenhum despoletou qualquer anomalia comportamental, uma vez que para o caso de uso U1 o comportamento aprendido é igual ao comportamento atual, enquanto que para os casos de uso U2 e U3 o comportamento aprendido é sempre igual ou superior ao comportamento atual.

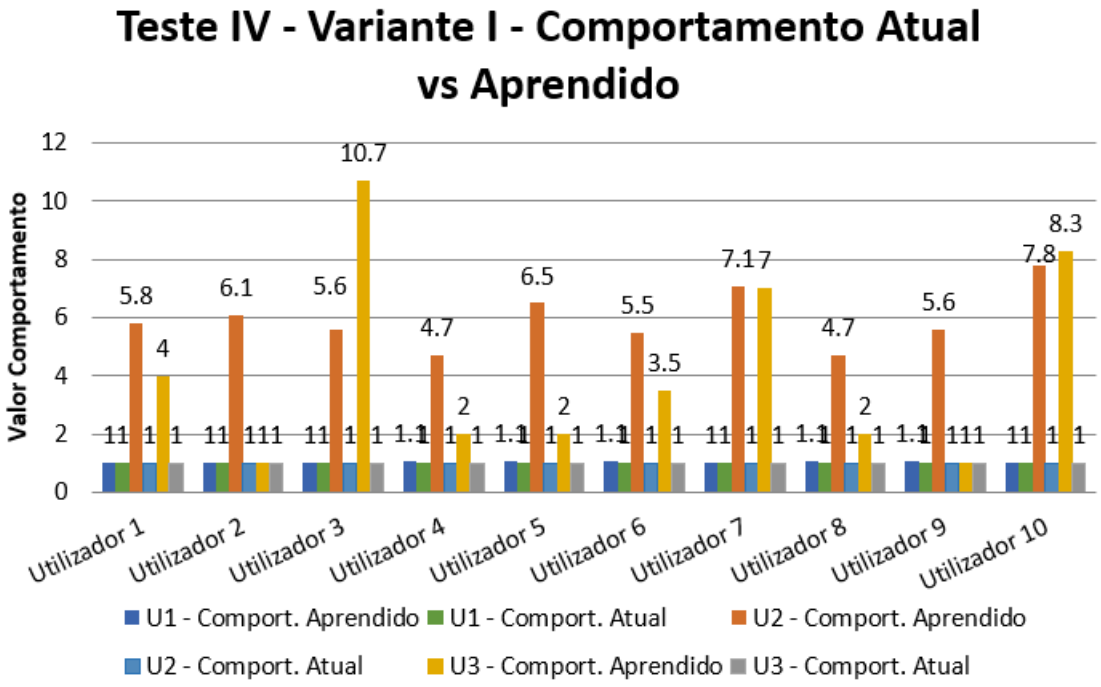

Figura 5.10: Teste IV - Variante I - Comportamento Atual vs Aprendido

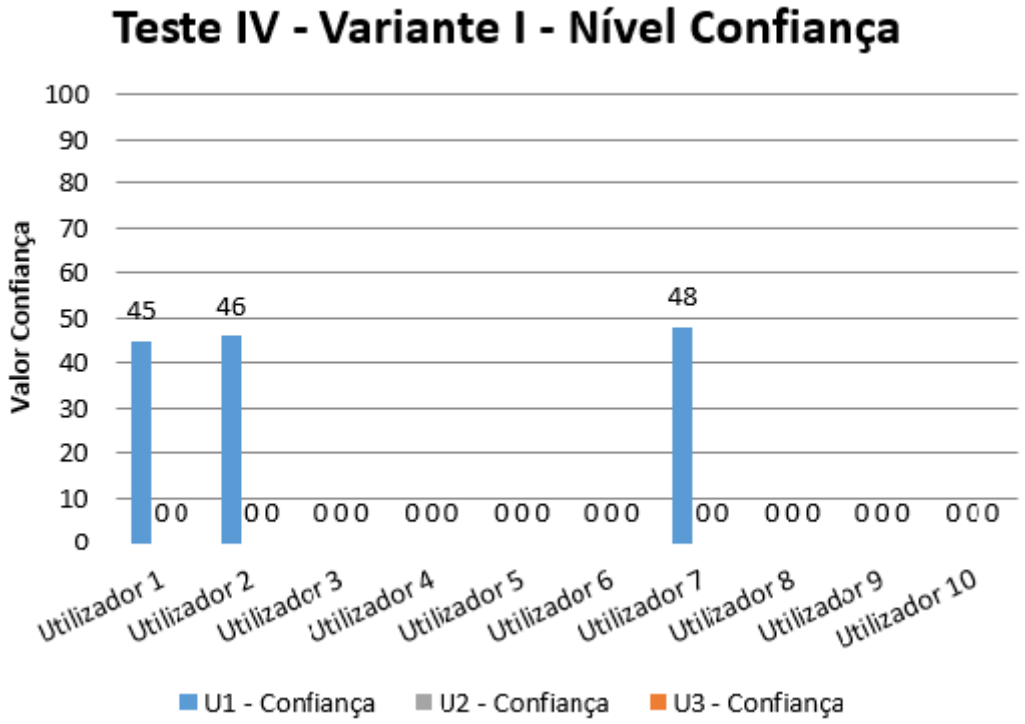

Figura 5.11: Teste IV - Variante I - Nível de Confiança

Com a realização da segunda variante deste teste, foi possível comprovar que o comportamento atual de dois utilizadores deu origem à geração de uma anomalia, ambas com uma confiança de 100%. Para os restantes utilizadores, apesar de o comportamento atual e aprendido ser idêntico ao dos utilizadores anteriores, não foi despoletada qualquer anomalia, permanecendo mesmo uma confiança de 0% para todos eles. Quanto aos casos de uso U2 e U3, constata-se que o comportamento atual é muito semelhante ao comportamento aprendido para a generalidade dos utilizadores, pelo que não foi registada qualquer anomalia comportamental. Contudo, importa referir que se considera os valores de comportamento aprendido como elevados. Os resultados desta variante são os ilustrados nas figuras 5.12 e 5.13.

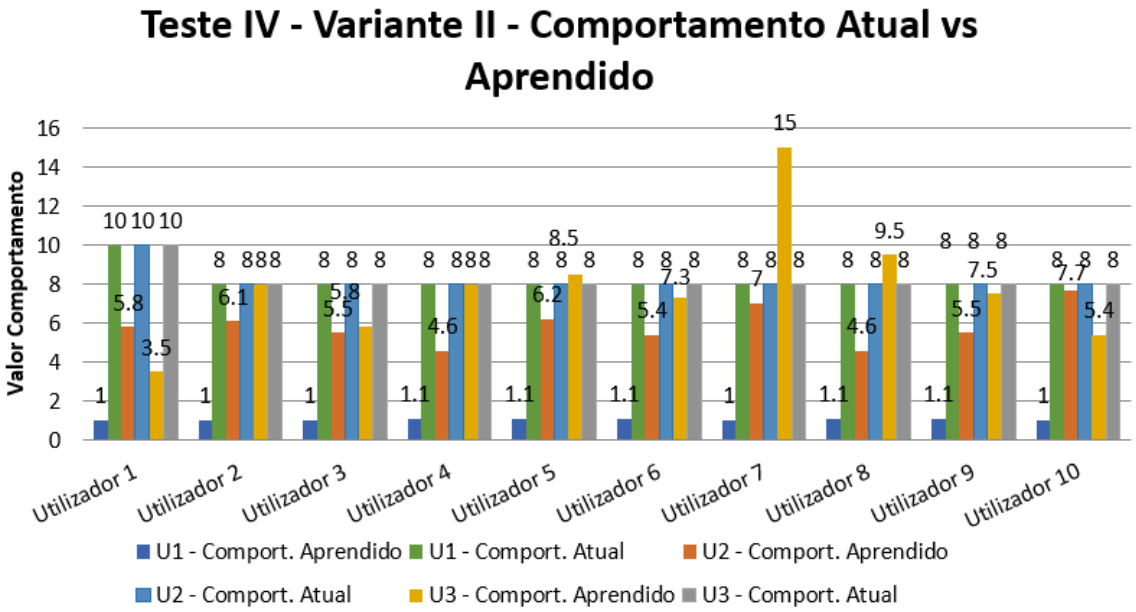

Figura 5.12: Teste IV - Variante II - Comportamento Atual vs Aprendido

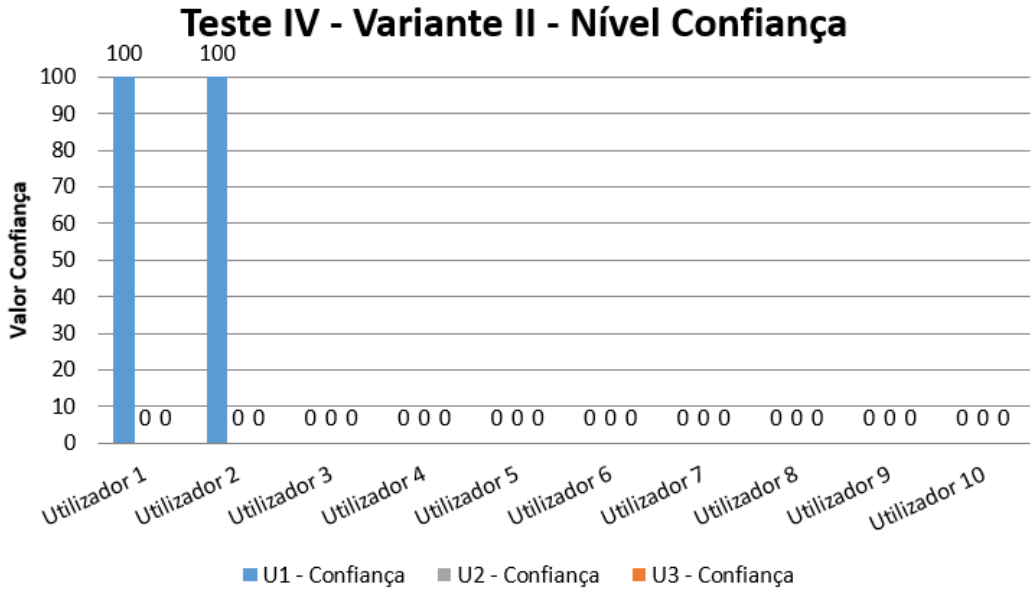

Figura 5.13: Teste IV - Variante II - Nível de Confiança

5.2 Análise dos Resultados

Concluída a apresentação dos resultados obtidos nos diferentes testes realizados, importa analisá-los, discuti-los e compará-los de uma forma global e abrangente. Começando, em primeiro lugar, por analisar as diferentes variantes de estudo, conforme definidas na tabela 4.2, chega-se à conclusão que tanto o intervalo de confiança, como o período de retenção de dados não apresentaram, nos resultados obtidos, quaisquer diferenças nos modelos computados para as diferentes variantes em estudo. Teoricamente, quanto maior for o período de retenção de dados, mais lentamente variará a distribuição, e mais estável será o modelo, gerando menos falsos positivos. Contudo, nos casos de uso estudados não foi observada qualquer diferença causada por este parâmetro. Quanto ao intervalo de confiança, a geração de anomalias foi sempre provocada por comportamentos aos quais foi atribuída uma confiança de 100%, fator este que levou todas as variantes a apresentarem um comportamento igual. Neste sentido, e desconhecendo-se a forma como este período de retenção contribui para a construção dos modelos (uma vez que essa informação é proprietária da *IBM* e não foi disponibilizada pela mesma), conclui-se que devem ser utilizados como parâmetros de referência um intervalo de confiança de 95%, e um período de retenção de dados de trinta dias, valores estes disponibilizados, por omissão, pela aplicação na criação dos modelos.

Previamente à análise dos quatro testes realizados, importa referir que o utilizador identificado nos resultados como *Utilizador 1* corresponde ao utilizador para o qual foi feito uso dos dois primeiros *scripts* identificados no Capítulo 4, Secção 4.4.

O primeiro teste realizado viu o seu objetivo plenamente cumprido, pois permitiu confirmar que os modelos criados minimizam a criação de falsos positivos para os casos de uso U1 e U2, não tendo sido, inclusivamente, detetado qualquer caso nos testes realizados. O caso de uso U3, quando sujeito a valores anormais de autenticação, levou à produção de alertas de anomalias comportamentais, conforme esperado. Assim, os resultados obtidos neste teste confirmam as expectativas iniciais. Contudo, para o caso de uso U1, se um dado utilizador fizer uso de uma estação de trabalho diferente da sua habitual, durante o seu período normal de trabalho, esta situação acabará por não ser detetada por este modelo. Por outro lado, a partir deste primeiro teste é possível tirar uma outra conclusão, a qual é comum e transversal a todos os outros testes realizados. A aplicação ML, quer para os modelos genéricos por omissão, quer para os modelos personalizados, não permite deteções em tempo-real ou perto de tempo-real, uma vez que os eventos ocorridos numa dada hora só são processados no final dessa mesma hora, ou seja, ela faz uso de intervalos de uma hora para processamento dos dados de eventos gerados, não sendo esse mesmo intervalo configurável, o que leva a que uma dada anomalia, no pior dos casos, possa ser reportada até uma hora após a sua ocorrência/registo.

Na sequência do segundo teste verificou-se que os resultados foram consistentes para ambas as variantes, tendo o modelo correspondente ao caso de uso U1 funcionado como previsto. Para mais, ele despoletou anomalia apenas para um dos utilizadores, uma vez que a confiança para os restantes utilizadores permaneceu sempre no valor zero, isto apesar de também detetar uma alteração substancial de comportamento nesses mesmos utilizadores. Tal facto pode dever-se a uma maior necessidade de dados de treino para os modelos, de modo a apresentarem um comportamento mais confiável e consistente. No entanto, garantindo essa consistência, considera-se que o modelo definido permite a deteção de situações comportamentais anómalas a partir da utilização de pelo menos uma estação de trabalho distinta da estação habitual. Quanto ao caso de uso U3, este apresentou também o comportamento esperado, não originando quaisquer anomalias, uma vez que este teste não visava uma alteração substancial ao comportamento normal dos utilizadores.

O terceiro teste, no caso da primeira variante, mostrou que os utilizadores apresentavam um comportamento atual sempre igual ou inferior ao aprendido, para ambos os casos de uso U2 e U3. Já na

segunda variante, com o aumento do número de tentativas de autenticação, o comportamento atual subiu, ficando acima do comportamento aprendido. No entanto, não foi possível constatar a geração de quaisquer anomalias para ambas as variantes, apresentando sempre um nível de confiança de 0% para todos os utilizadores. A existência de valores de comportamento aprendido tão elevados, para ambos os casos de uso, não era expectável, isto é, considerando que este teste decorre no período não habitual de trabalho, seria de esperar que os valores de comportamento aprendido fossem na ordem de zero ou muito próximo desse valor, não sendo contudo o que se observa. Neste sentido, e analisando na aplicação os gráficos comportamentais dos diferentes utilizadores, percebe-se que é feita uma interpolação entre os dois últimos pontos com registo de atividade para esse utilizador, o que, de certo, não traduzirá o comportamento real e correto, originando desta forma resultados incorretos e não esperados. Deste modo, comprova-se que os modelos apresentam um comportamento diferente do preconizado, seja na documentação do fabricante, seja na própria aplicação, distorcendo assim os resultados e não tornando possível considerar os modelos associados ao caso de uso U2 e U3 como válidos.

O quarto e último teste, para que conseguisse cumprir o objetivo proposto, necessitava que fossem despoletadas anomalias, simultaneamente, pelos casos de uso U1 e U2. Não obstante este facto, a análise destes pode ser realizada de forma separada. Para o caso de uso U1, verifica-se que na primeira variante não foi despoletada nenhuma anomalia, enquanto que na segunda houve dois utilizadores que apresentaram um comportamento anómalo. Contudo, uma vez mais, observa-se que o comportamento aprendido apresenta valores não iguais a zero, o que impede a geração de anomalias nesta primeira variante. Por outro lado, importa dizer também que apenas dois utilizadores geram anomalias, isto quando possuem um tempo de treino e comportamento aprendido semelhante a outros utilizadores testados. Quanto aos casos de uso U2 e U3, repete-se o que já tinha acontecido no terceiro teste, pelo que os comentários são idênticos. Assim, conclui-se também que este teste não demonstrou resultados de acordo com o esperado, por fatores associados ao desempenho da aplicação em análise, e que ultrapassam o âmbito deste trabalho.

Após esta análise aos quatro testes realizados, é possível sumarizar na tabela 5.1 as principais vantagens e desvantagens que decorrem da utilização dos modelos genéricos disponibilizados por omissão, em contraste com a utilização dos modelos personalizados que podem ser criados pelos analistas.

Tabela 5.1: Vantagens e Desvantagens dos Modelos por Omissão vs Modelos Personalizados

	Modelos por Omissão	**Modelos Personalizados**
Vantagens	- Prontos a utilizar, sem necessidade de parametrização - Parâmetros configurados por omissão apresentam um bom comportamento - Modelos abrangentes divididos por âmbitos, nomeadamente pelas diferentes categorias de alto nível de eventos	- Capacidade de personalização, conforme necessidade da organização
Desvantagens	- Desconhece-se o modo como os algoritmos de aprendizagem automática utilizados são implementados - Visão e monitorização completa requer a ativação de grande parte destes modelos - Deteção não é feita em tempo-real - Código fonte não público	- Desconhece-se o modo como os algoritmos de aprendizagem automática utilizados são implementados - Visão parcial e particular, dependendo do caso de uso implementado - Deteção não é feita em tempo-real - Código fonte não público - Possibilidades de parametrização reduzidas e limitadas

Para finalizar esta análise, importa também referir um conjunto de factores que possuem um impacto direto nos resultados obtidos, e que, por esse motivo, não devem ser deixados de parte:

- É imprescindível garantir que o *QRadar* consegue interpretar, descodificar e traduzir os diferentes tipos de *logs* recebidos para o seu formato interno, devendo ser sempre validada esta transformação, por forma a garantir a melhor correção dos conjuntos de dados de treino;
- A aplicação ML possui um limite máximo de utilizadores que podem ser seguidos pelos diferentes modelos analíticos. Como tal, todos os utilizadores identificados como sendo alvo de seguimento por estes modelos, devem ver ativada a opção *Always Track with ML* na sua página de detalhes (aumentando, assim, a prioridade de seguimento deste utilizador). Caso possível é desejável colocar esse utilizador numa *watchlist*, definindo também a prioridade de seguimento, conforme descrito no Capítulo 4, Secção 4.4;
- Por último, constatou-se também que a existência de teletrabalho leva ao incremento da utilização de VPN's. Contudo, sempre que um dado utilizador não esteja ligado à mesma, não são recebidos quaisquer eventos que ocorram na sua estação de trabalho, ou só serão posteriormente sincronizados com o SIEM, o que leva a uma receção de eventos e monitorização diferente de tempo-real. Tal poder-se-á traduzir consequentemente em erros, falsos positivos e/ou atraso na deteção de situações anómalas.

Capítulo 6

Conclusões e Trabalho Futuro

6.1 Conclusões

O principal objetivo deste projeto passava por validar a exequibilidade de implementação de uma solução, de forma eficaz e eficiente, para deteção e análise de anomalias comportamentais de segurança dos utilizadores de uma grande organização. Este objetivo foi atingido com sucesso, tendo-se concluído que, com a solução avaliada, não é exequível detetar e analisar anomalias comportamentais de segurança dos utilizadores de forma eficaz e eficiente.

Da análise efetuada conclui-se que a ferramenta utilizada não apresentou o comportamento desejado e esperado, perante os diferentes casos de teste, pelo que se considera que a mesma não apresenta um bom grau de precisão, nem fornece alertas em tempo quase real, havendo assim certamente um caminho a percorrer para melhorar este desempenho. Para chegar a estas conclusões, foi desenvolvido este trabalho, sendo apresentado de seguida um resumo do mesmo, bem como as limitações a que este projeto esteve sujeito.

Após uma análise inicial do problema apresentado pela Altice Portugal, foram identificadas as principais preocupações de monitorização ao nível de utilizadores e atividades de interesse. Neste seguimento, foi simultaneamente definido um conjunto de casos de uso, bem como identificadas as fontes que contribuem com os dados necessários para essa mesma monitorização.

Posteriormente, foi elaborado um estudo exaustivo das aplicações UBA e ML, explorando todas as suas funcionalidades e compreendendo o seu funcionamento. Adicionalmente, foram também implementados e testados modelos analíticos personalizados, correspondentes aos casos de uso mapeados, bem como estudados quais os valores dos parâmetros desses modelos que permitiam a estes apresentar um melhor comportamento e capacidade de deteção. Mais, decidiu-se ir mais além, ativando um modelo genérico de atividade de autenticação, por forma a viabilizar a comparação entre os dois tipos de modelos disponibilizados, salientando assim as vantagens e desvantagens de cada um.

Este projeto, contudo, esteve sujeito a algumas limitações. A aplicação ML, utilizada para desenho e implementação dos modelos analíticos, é uma aplicação de uso comercial, cujo código fonte não está disponível publicamente, não tornando possível conhecer com maior detalhe quais os algoritmos de aprendizagem automática utilizados e a forma como os mesmos foram implementados. Para mais, tal facto inviabiliza também que seja feita alguma afinação ou ajuste, de modo a melhorar o desempenho dos modelos. Por outro lado, e apesar desta permitir aos analistas a criação de modelos personalizados, as possibilidades de configuração/parametrização destes são reduzidas e limitadas. Na prática, a aplicação ML deixa transparecer a ideia de que a utilização dos modelos genéricos disponibilizados por omissão deve ser a regra, dando também a possibilidade de criação de modelos personalizados, sem que este seja o principal foco desta.

Em conclusão, a aplicação ML não provou ser capaz de detetar anomalias comportamentais de segurança dos utilizadores para os diferentes casos de uso implementados. Paralelamente, permitiu deduzir que existem mais vantagens de utilização dos modelos genéricos disponibilizados, mesmo fazendo uso dos valores configurados por omissão, em contraste com os modelos personalizados. Para finalizar, este projeto demonstrou também a forma como é possível criar e definir uma metodologia que permita caraterizar as necessidades de monitorização específicas de uma grande organização, mais em particular na identificação, especificação e priorização de uma lista de casos de uso, bem como na escolha e identificação das fontes de dados necessárias.

6.2 Trabalho Futuro

No decurso deste projeto foram identificadas algumas oportunidades de trabalho futuro, as quais visam contribuir não só para a qualidade do produto final implementado, como também realizar uma análise mais global e transversal.

A primeira oportunidade de trabalho futuro passa pelo necessário enriquecimento do SIEM *QRadar* com eventos de outros tipos (não apenas de autenticação), originados tanto em estações de trabalho, mas sobretudo servidores, ou outras fontes que venham a ser identificadas. Simultaneamente, deverão ser ativados os diferentes tipos de modelos existentes na aplicação ML procurando assim detetar desvios comportamentais dos utilizadores em relação a si próprios ou aos seus grupos de pares, de uma forma transversal, através da monitorização e contribuição de diferentes categorias de eventos de alto nível. Esta oportunidade de trabalho futuro procura, assim, não se focar apenas em casos de uso particulares, mas sim fazer uso de uma visão mais global, tirando um proveito mais efetivo dos diferentes modelos de *Machine Leaning* genéricos e por omissão existentes na aplicação, e da existência de um perfil de risco por utilizador.

A segunda oportunidade de trabalho futuro resulta da vantajosa comparação entre os resultados obtidos pela utilização dos modelos analíticos disponibilizados pela aplicação UBA e o recurso a técnicas comuns de *Machine Leaning*, supervisionadas ou não, fora desta, tendo como base a utilização dos mesmos eventos para o mesmo período de tempo. Esta comparação permitirá compreender a real eficácia dos modelos desta aplicação, bem como estudar quais os parâmetros que mais influenciam esses mesmos modelos.

Por fim é relevante dizer que este projeto não deve ser encarado como finito no tempo, mas sim como um processo contínuo de constante desenvolvimento e aperfeiçoamento, sendo recomendável a sua revisão e avaliação regular, de forma a garantir uma capacidade de deteção de comportamentos anómalos permanente e atualizada, de acordo com as ameaças emergentes.

Bibliografia

[1] T. Matthews, "Insider threats: How to stop the most common and damaging security risk you face." Disponível em: https://www.exabeam.com/ueba/insider-threats/. Acedido em: 11 de novembro de 2019.

[2] V. Vasudevan, "Use cases for security analytics." Disponível em: https://informationsecurity.report/Resources/Whitepapers/0bb5fb53-e36e-41dc-957c-b32e38453d73_UseCasesforSecurityAnalytics.pdf. Acedido em: 17 de outubro de 2019.

[3] G. Sadowski, A. Litan, T. Bussa, , and T. Phillips, "Market guide for user and entity behavior analytics," tech. rep., Gartner, Inc., Abril 2018.

[4] Aruba, "The ciso's guide to machine learning & user and entity behavioral analytics." Disponível em: https://www.arubanetworks.com/assets/CisoGuide.pdf. Acedido em: 23 de setembro de 2019.

[5] K. Kavanagh, T. Bussa, and G. Sadowski, "Magic quadrant for security information and event management," tech. rep., Gartner, Inc., Fevereiro 2020.

[6] IBM, "Qradar architecture overview." Disponível em: https://www.ibm.com/support/knowledgecenter/en/SS42VS_7.3.2/com.ibm.qradar.doc/c_qradar_deployment_guide_arch.html. Acedido em: 20 de novembro de 2019.

[7] IBM, "Ibm qradar user behavior analytics (uba) app version 3.4.0 - user guide," manual, IBM Corporation, 2019.

[8] S. Bhatt, P. K. Manadhata, and L. Zomlot, "The operational role of security information and event management systems," *IEEE Security Privacy*, vol. 12, pp. 35–41, Setembro 2014.

[9] A. Osório, "Threat detection in siem considering risk assessment," Master's thesis, Universidade de Lisboa - Faculdade de Ciências, 2018.

[10] C. Allen, K. H. Pherson, D. Thomas, V. Corsi, M. Gardiner, S. MacIsaac, D. McGarvey, and R. Thompson, "Assessing the mind mind of the malicious insider: using a behavioral model and data analytics to improve continuous evaluation," tech. rep., INSA, Abril 2017.

[11] X. Xiangyu, T. Zhang, D. Du, G. Zhao, Q. Gao, W. Zhao, and S. Zhang, "Method and system for detecting anomalous user behaviors: An ensemble approach," in *The 30th International Conference on Software Engineering and Knowledge Engineering*, pp. 263–307, Julho 2018.

[12] L. Liu, O. De Vel, Q. Han, J. Zhang, and Y. Xiang, "Detecting and preventing cyber insider threats: A survey," *IEEE Communications Surveys Tutorials*, vol. 20, pp. 1397–1417, Fevereiro 2018.

[13] H. Schulze, "Insider threat 2018 report," tech. rep., Cibersecurity Insiders, 2018.

[14] IBM, "An integrated approach to insider threat protection." Disponível em: `https://www.ibm.com/downloads/cas/GRQQYQBJ.pdf`. Acedido em: 17 de setembro de 2019.

[15] D. Shackleford, "Using analytics to predict future attacks and breaches," tech. rep., SANS Institute, Janeiro 2016.

[16] A. Chuvakin and A. Barros, "Demystifying security analytics: Sources, methods and use cases," tech. rep., Gartner, Inc., Março 2017.

[17] A. Pritz, *Security Analytics*, pp. 5–9. John Wiley & Sons, Inc., 2018.

[18] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, Julho 2009.

[19] W. Ma, *User Behavior Pattern Based Security Provisioning for Distributed Systems*. PhD thesis, University of Ontario Institute of Technology, Oshawa, Ontario, Canada, 2016.

[20] Balabit, "The essential guide to user behavior analytics." Disponível em: `https://www.ciosummits.com/Online_Assets_Balabit_Essential_Guide_to_User_Behavior_Analytics.pdf`. Acedido em: 3 de outubro de 2019.

[21] M. Shashanka, M. Shen, and J. Wang, "User and entity behavior analytics for enterprise security," in *2016 IEEE International Conference on Big Data (Big Data)*, pp. 1867–1874, Dezembro 2016.

[22] M. B. Salem and S. J. Stolfo, "Masquerade attack detection using a search-behavior modeling approach," in *Columbia University Computer Science Technical Reports*, (New York, USA), Department of Computer Science, Columbia University, Julho 2010.

[23] S. Angeletou, M. Rowe, and H. Alani, "Modelling and analysis of user behaviour in online communities," in *The Semantic Web – ISWC 2011* (L. Aroyo, C. Welty, H. Alani, J. Taylor, A. Bernstein, L. Kagal, N. Noy, and E. Blomqvist, eds.), (Berlin, Heidelberg), pp. 35–50, Springer Berlin Heidelberg, 2011.

[24] Q. Hu, B. Tang, and D. Lin, "Anomalous user activity detection in enterprise multi-source logs," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 797–803, Novembro 2017.

[25] P. Thompson, "Weak models for insider threat detection," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5403, Setembro 2004.

[26] M. A. Maloof and G. D. Stephens, "elicit: A system for detecting insiders who violate need-to-know," in *Recent Advances in Intrusion Detection* (C. Kruegel, R. Lippmann, and A. Clark, eds.), (Berlin, Heidelberg), pp. 146–166, Springer Berlin Heidelberg, 2007.

[27] H. Eldardiry, E. Bart, J. Liu, J. Hanley, B. Price, and O. Brdiczka, "Multi-domain information fusion for insider threat detection," in *2013 IEEE Security and Privacy Workshops*, pp. 45–51, Maio 2013.

[28] IBM, "User behavior analytics rule." Disponível em: `https://www.ibm.com/support/knowledgecenter/SS42VS_7.3.2/com.ibm.UBAapp.doc/c_Qapps_UBA_intro.html`. Acedido em: 8 de outubro de 2019.

[29] T. Ryan, "Demystifying machine learning and anomaly detection: Practical applications in splunk for insider threat detection and network security analytics." Disponível em: http://splk.it/2tjtKLU, 2016. Acedido em: 9 de outubro de 2019.

[30] A. Lashkari, M. Chen, and A. Ghorbani, "A survey on user profiling model foranomaly detection in cyberspace," *Journal of Cyber Security and Mobility*, vol. 8, pp. 75–112, Outubro 2018.

[31] B. Ware, R. Kerr, C. Knisley, E. Wright, and T. Read, "User behavior analytics, redefined." Disponível em: https://haystax.com/wp-content/uploads/2017/05/Haystax-UBA-Redefined-White-Paper_DIGITAL_FINAL.pdf, 2017. Acedido em: 8 de outubro de 2019.

[32] A. Chuvakin and A. Barros, "How to develop and maintain security monitoring use cases," tech. rep., Gartner, Inc., Janeiro 2018.

[33] Exabeam, "The essential guide to siem." Disponível em: https://www.exabeam.com/siem-guide/ueba/. Acedido em: 7 de novembro de 2019.

[34] D. Shackleford, "Sans 2016 security analytics survey," tech. rep., SANS Institute, Dezembro 2016.

[35] R. Santos, "Controlos de cibersegurança em ambientes ms windows de grandes empresas: Integração efetiva de eventos relevantes de segurança no siem alienvault usm," Master's thesis, Universidade de Lisboa - Faculdade de Ciências, 2018.

[36] O. Cassetto, "What ueba stands for." Disponível em: https://www.exabeam.com/ueba/what-ueba-stands-for/. Acedido em: 6 de novembro de 2019.

[37] O. Cassetto, "User behavior analytics (uba/ueba): The key to uncovering insider and unknown security threats." Disponível em: https://www.exabeam.com/ueba/user-behavior-analytics/. Acedido em: 6 de novembro de 2019.

[38] Splunk, "What is user behavior analytics (uba)/user entity behavior analytics (ueba)?." Disponível em: https://www.splunk.com/en_us/data-insider/user-behavior-analytics-ueba.html. Acedido em: 6 de novembro de 2019.

[39] L. Voigt, "What is ueba and why it should be an essential part of your incident response." Disponível em: https://www.exabeam.com/ueba/ueba-uba-siem-incident-response/. Acedido em: 7 de novembro de 2019.

[40] IBM, "Qradar events and flows." Disponível em: https://www.ibm.com/support/knowledgecenter/en/SS42VS_7.3.2/com.ibm.qradar.doc/c_qradar_deploy_event_and_flow_pipeline.html. Acedido em: 20 de novembro de 2019.

[41] IBMCorporation, "Beat insider threats with integrated user behavior analytics." Disponível em: https://www.midlandinfosys.com/pdf/qradar-siem-security-machine-learning-internal-threats.pdf, Fevereiro 2018. Acedido em: 11 de setembro de 2019.

[42] IBM, *IBM QRadar User Behavior Analytics (UBA) app Version 3.4.0 - User Guide*. IBM Corporation, 2019.

Anexos

Anexo A

Instalação UBA App

O procedimento aqui descrito detalha os passos realizados para a instalação da aplicação UBA no sistema *IBM QRadar*. Este procedimento, contudo, não dispensa a leitura do manual de utilizador desta mesma aplicação [42]. Previamente à instalação, torna-se necessário garantir que alguns requisitos são cumpridos:

- Garantir que o utilizador do *QRadar* tem um perfil de administração, capaz de adicionar novas extensões;
- Existência 1.2 GB de memória livre disponível da *pool* de memória das aplicações;
- Garantir que o *IBM Sense DSM* se encontra instalado;
- Garantir a compatibilidade entre a versão do *IBM Security QRadar* e a UBA App. Neste caso, foi instalada a versão 3.4.0 da UBA App, compatível com a versão 7.3.1 do *IBM Security QRadar*.

Para proceder à instalação da UBA App, sem fazer uso do *QRadar Assistant App*, deve descarregar-se um arquivo, em formato comprimido, a partir da página web da *IBM Security App Exchange*[19] para a máquina local. Apesar de esta ser disponibilizada de forma gratuita, é indispensável que exista uma conta IBM para realizar esta ação.

Após esta etapa, o utilizador, acedendo ao menu de navegação, no separador *Admin*, deverá abrir *System Configuration* > *Extensions Management*, clicando em adicionar nova extensão, conforme mostra a figura A.1.

[19] https://exchange.xforce.ibmcloud.com/hub

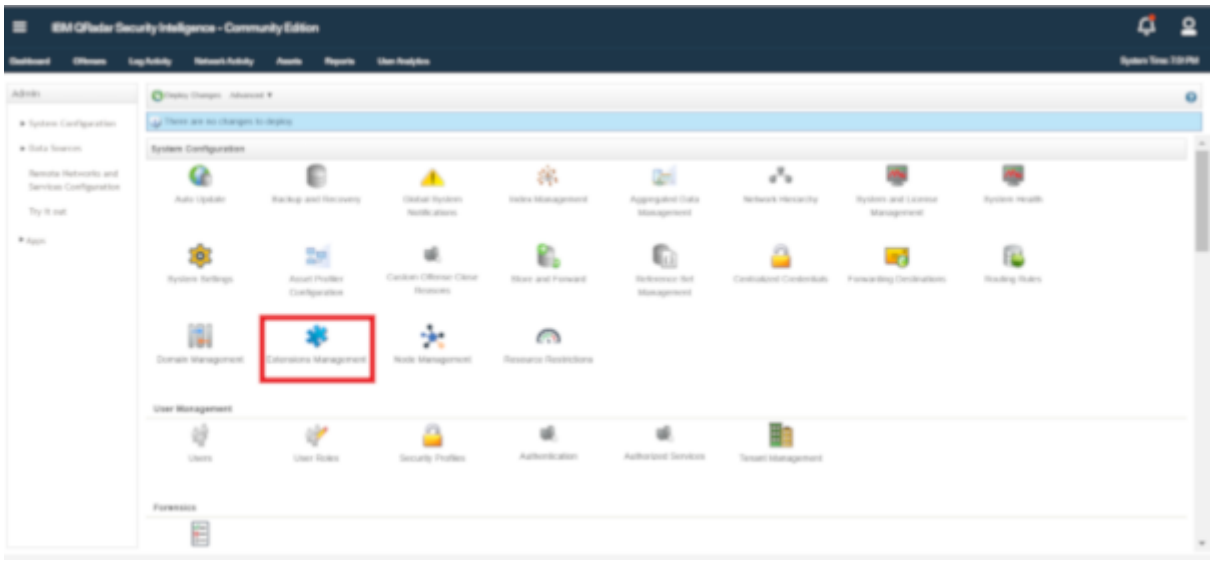

Figura A.1: Instalação UBA App - Passo 1: Configuração de Extensões

Ao executar esta ação, abrirá uma janela de acordo com a figura A.2.

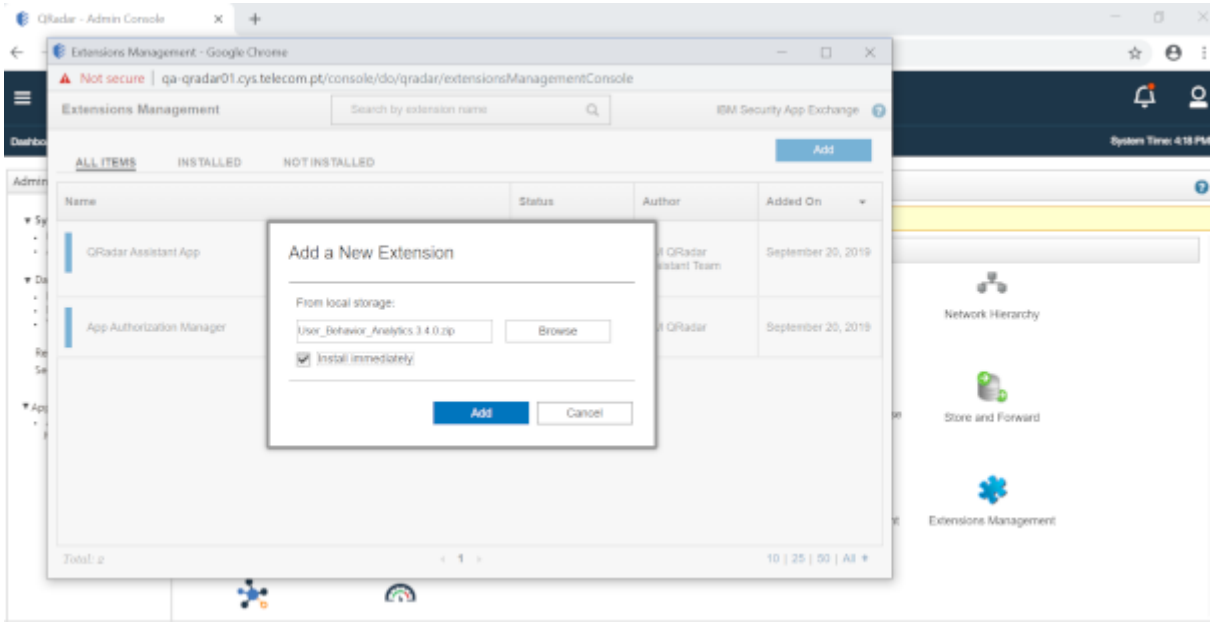

Figura A.2: Instalação UBA App - Passo 2: Adição de nova Extensão

Seguidamente, deverá ser selecionada a diretoria onde foi armazenado o arquivo descarregado, bem como selecionada a opção *Install Immediately*, clicando posteriormente em adicionar. Após isto, será necessário esperar alguns minutos até a aplicação ficar ativa, sendo que a instalação de alguns pacotes continuará a decorrer em segundo plano. O resultado final é visualizado na figura A.3

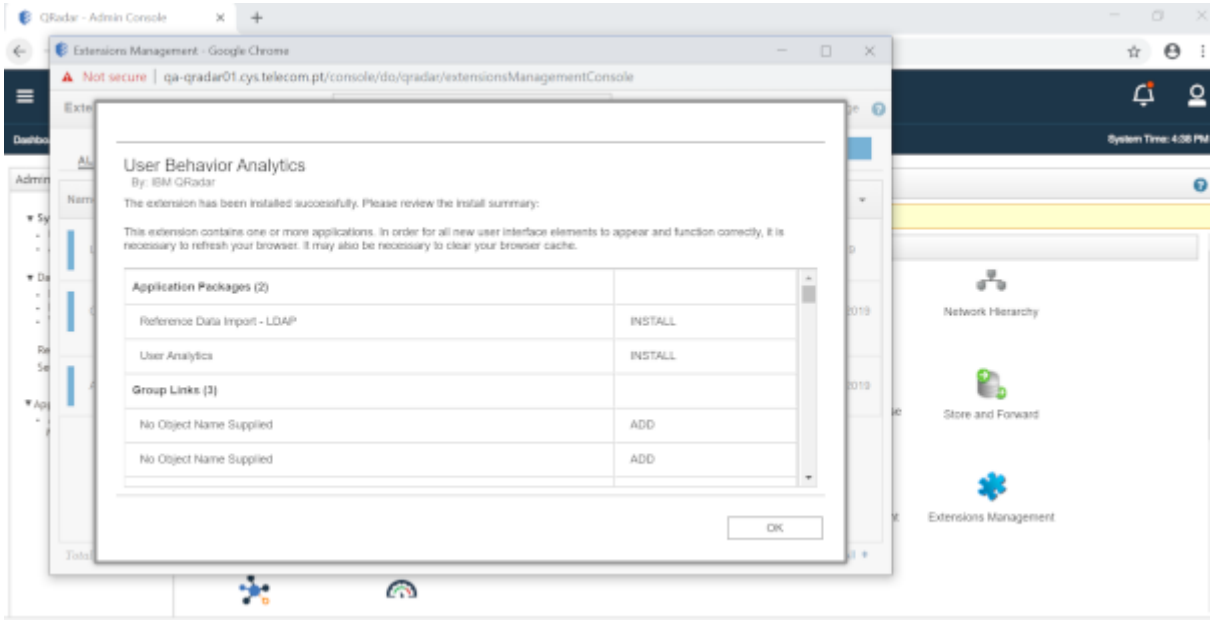

Figura A.3: Instalação UBA App - Passo 3: Resultado instalação da extensão

Concluída a instalação, é essencial executar um *Deploy Full Configuration*, bem como limpar a cache do navegador utilizado e refrescar a janela.

Anexo B

Configuração UBA App

Para configurar a aplicação UBA devem ser seguidos os seguintes passos:

1. Configuração do *Authorization Token*

    (a) No menu de navegação selecionar o separador *Admin*;

    (b) Clicar no ícone *UBA Settings (Apps > User Analytics > UBA Settings)*, aparecendo a imagem da figura B.1;

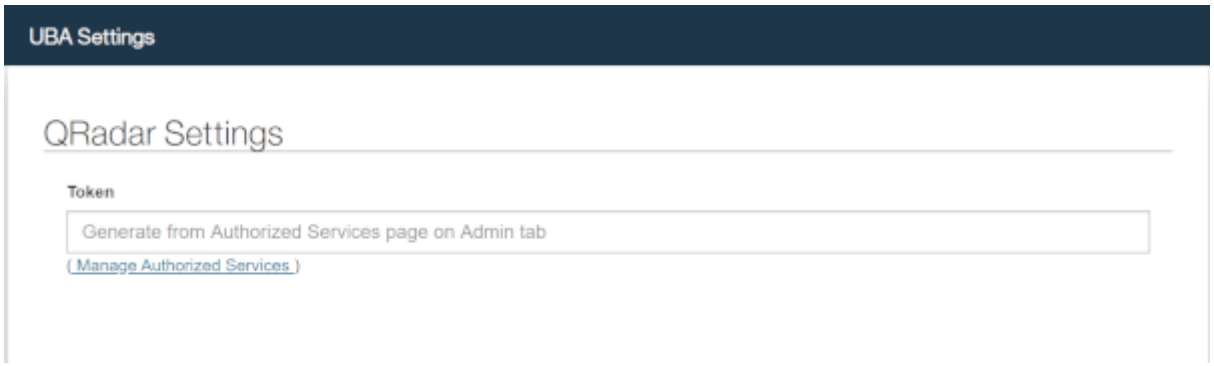

Figura B.1: Configuração do Authorization Token - Passo 1: UBA Settings

(c) Clicar na hiperligação *Manage Authorized Services*, a qual abrirá uma nova janela, clicando depois em *Add Authorized Service*. Aparecerá a janela para a adição do novo serviço, conforme a figura B.2;

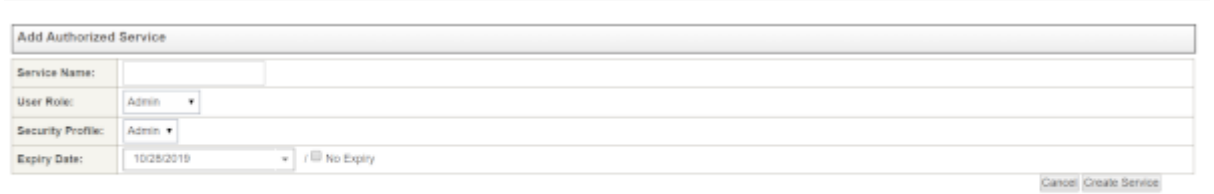

Figura B.2: Configuração do Authorization Token - Passo 2: Adicionar novo serviço

(d) Preencher o campo do nome do serviço com "UBA", selecionando o perfil *Admin*, nos campos *User Role* e *Security Profile*;

(e) Definir este serviço como não tendo data de expiração. No final, deverá aparecer preenchido como ilustrado na figura B.3;

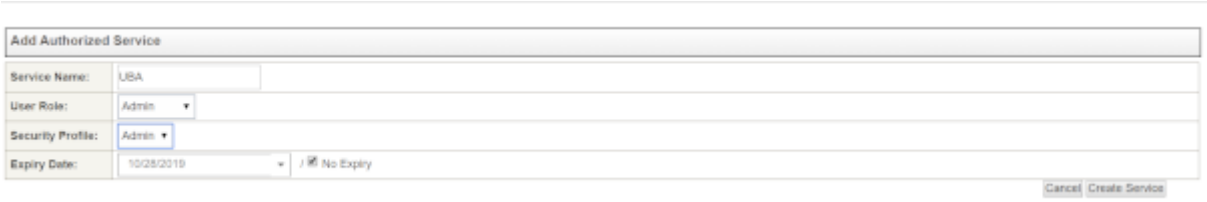

Figura B.3: Configuração do Authorization Token - Passo 3: Configuração do novo serviço

(f) Clicar em *Create Service*. Após esta ação o utilizador é devolvido ao menu anterior, conforme mostra a figura B.4;

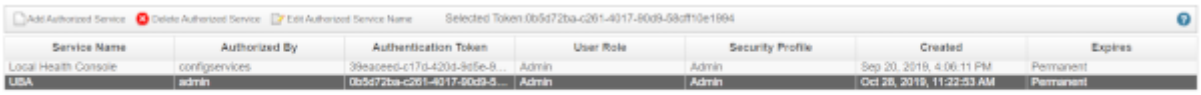

Figura B.4: Configuração Authorization Token - Passo 4: Criação do novo serviço

(g) Após esta ação, e apesar de não vir descrito no manual, é necessário realizar um *Deploy Full Configuration*;

(h) Quando concluído, retornar ao menu da figura B.4, selecionar o serviço "UBA"e copiar o *token* a partir do campo *Selected Token* na barra de menu;

(i) Regressando ao menu do *QRadar Settings* (figura B.1), introduzir o *token* anteriormente copiado no campo *Token*, conforme se observa na figura seguinte.

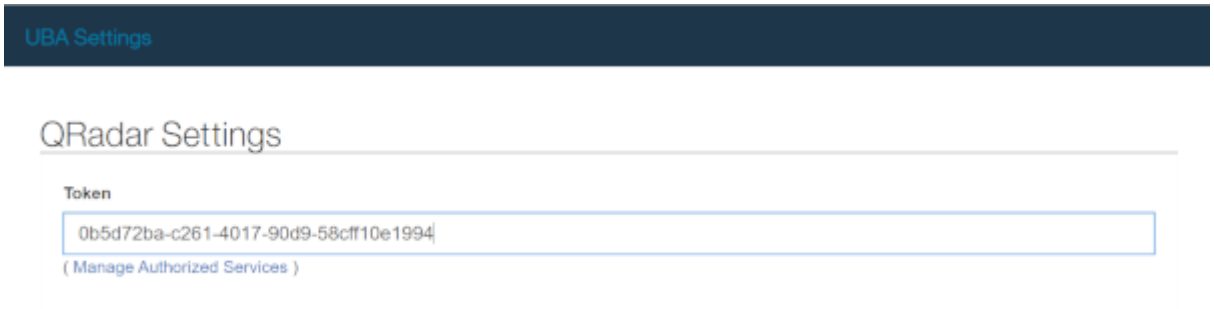

Figura B.5: Configuração do Authorization Token - Passo 5: Introdução do Token

2. Configuração de *Content Package Settings*

   (a) No menu de navegação selecionar o separador *Admin*;

   (b) Clicar no ícone *UBA Settings (Apps > User Analytics > UBA Settings)*;

   (c) Validar que na secção *Content Package Settings* a *checkbox* está ativa. Caso contrário, ativá-la, devendo aparecer como na imagem seguinte, figura B.6.

3. Configuração de *Application Settings*

   (a) No menu de navegação selecionar o separador *Admin*;

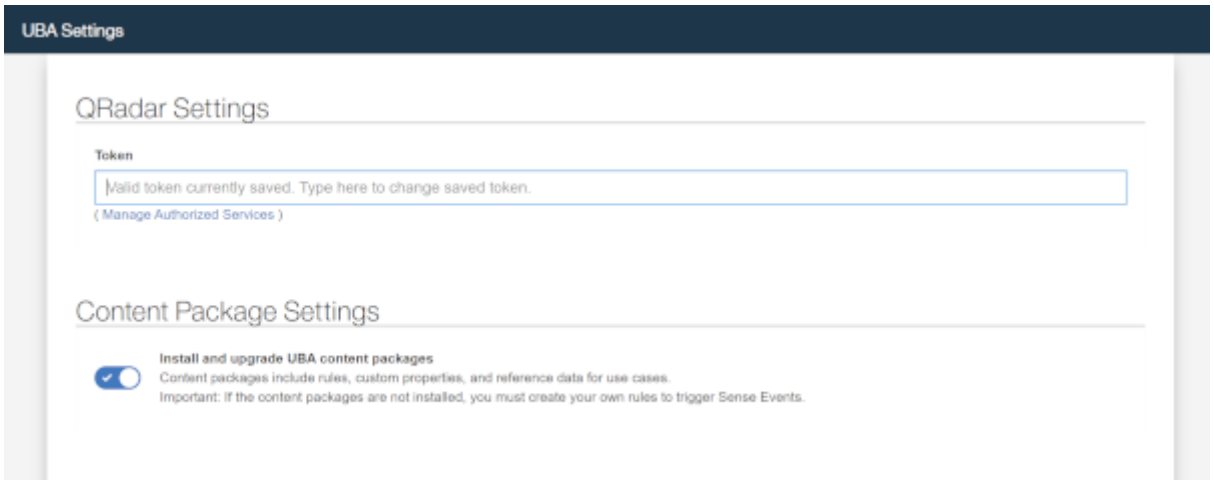

Figura B.6: Configuração de Content Package Settings

(b) Clicar no ícone *UBA Settings (Apps > User Analytics > UBA Settings)*;

(c) Na secção *Application Settings* configurar os seguintes parâmetros:

- *Risk Treshold* - seleccionar a opção *static* e definir o valor de *static risk treshold*. Simultaneamente, ativar a opção de gerar ofensas para os utilizadores de elevado risco;
- *Decay Risk by this factor per hour* - introduzir o valor de 0,05 neste campo.

Após estas configurações, obteve-se o resultado presente na figura B.7.

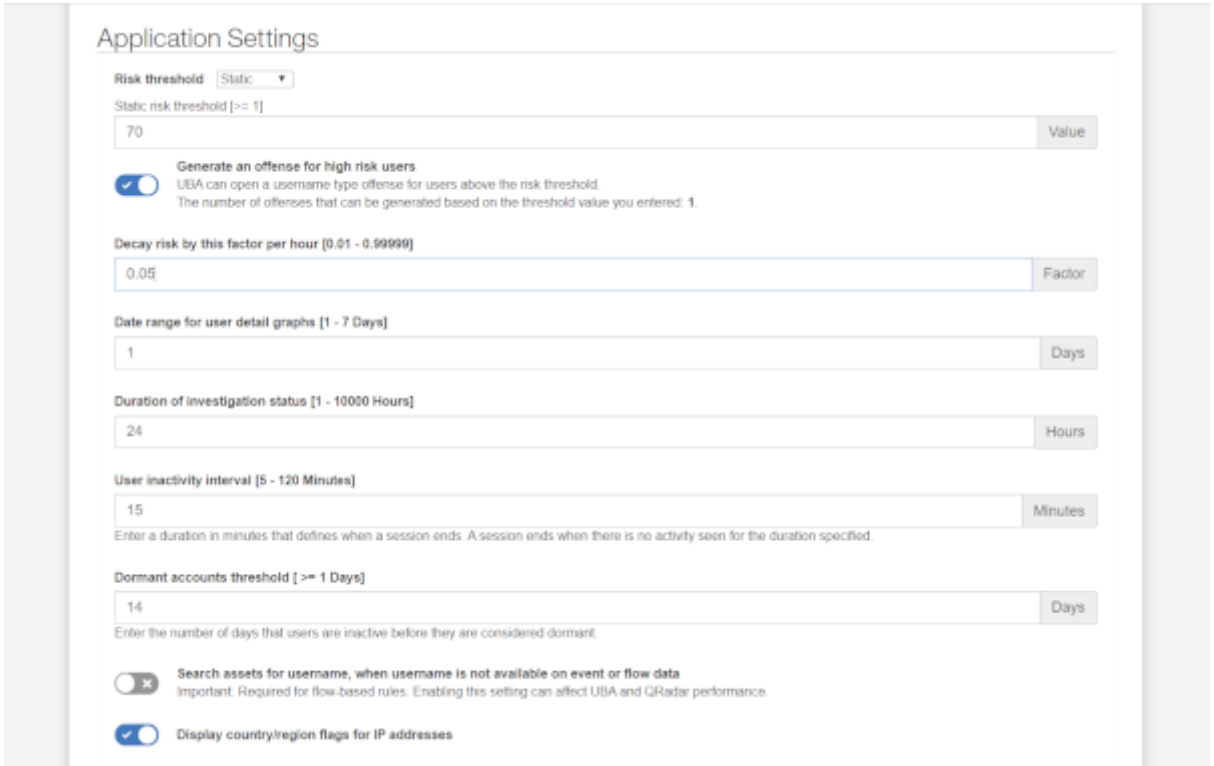

Figura B.7: Configuração de Application Settings

4. Ativação de Indexes

(a) No menu de navegação selecionar o separador *Admin*;

(b) Na secção *System Configuration*, clicar no ícone *Index Management*, conforme figura B.8;

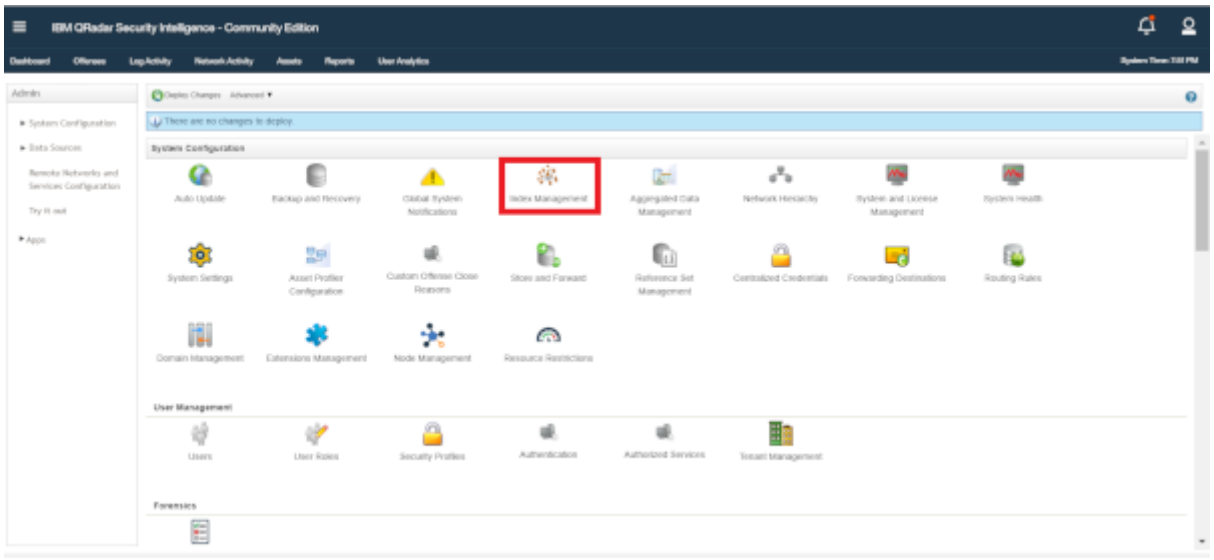

Figura B.8: Ativação de Indexes - Passo 1: Index Management

(c) Na página de gestão dos indexes, utilizar a caixa de pesquisa para pesquisar os seguintes indexes: *High Level Category*, *Low Level Category*, *Sense Value*, *Username* e *Sense Overall Score*, ativando aqueles que se encontrem inativos;

(d) No final, basta carregar em guardar, podendo observar-se uma disposição semelhante à da imagem abaixo.

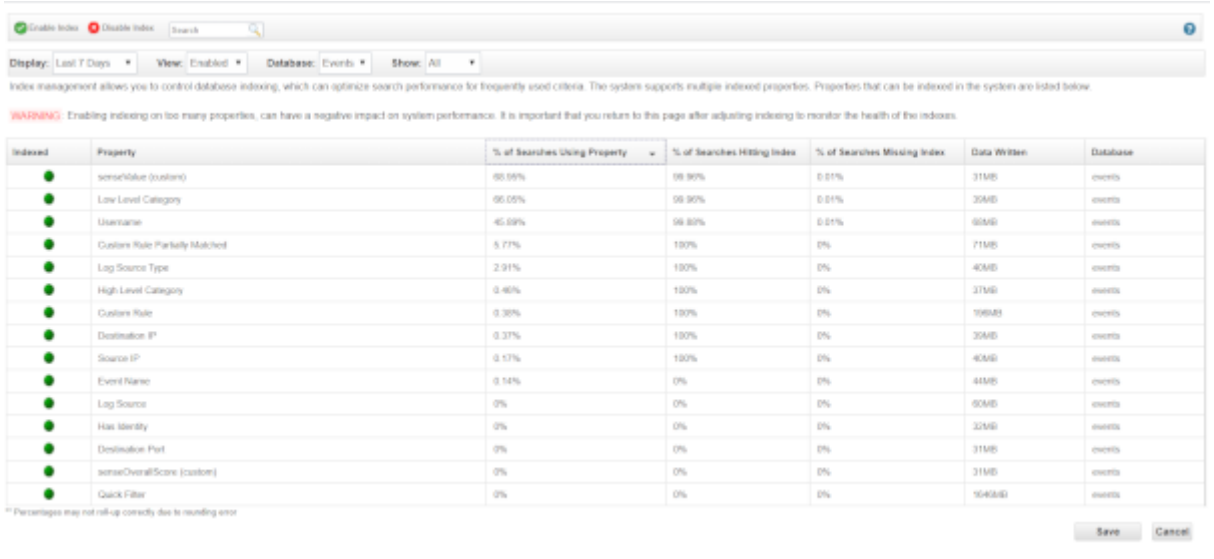

Figura B.9: Ativação de Indexes - Passo 2: Resultado final

Anexo C

Instalação ML App

À semelhança da instalação da UBA App, a instalação da *app* de *Machine Learning* também exige o cumprimento de alguns pré-requisitos. Esses pré-requisitos são:

- Existência de uma versão *IBM Security QRadar* igual ou superior à 7.3.1;
- Existência de 2 GB de memória livre disponível da *pool* de memória das aplicações;
- Garantia que a aplicação UBA está instalada, confirmando que no separador *User Analytics* existem dados.

Para proceder à instalação da ML App, clicar no ícone *Machine Learning Settings*, existente na secção *User Analytics* do separador *Admin*, o qual é acedido através do menu de navegação.

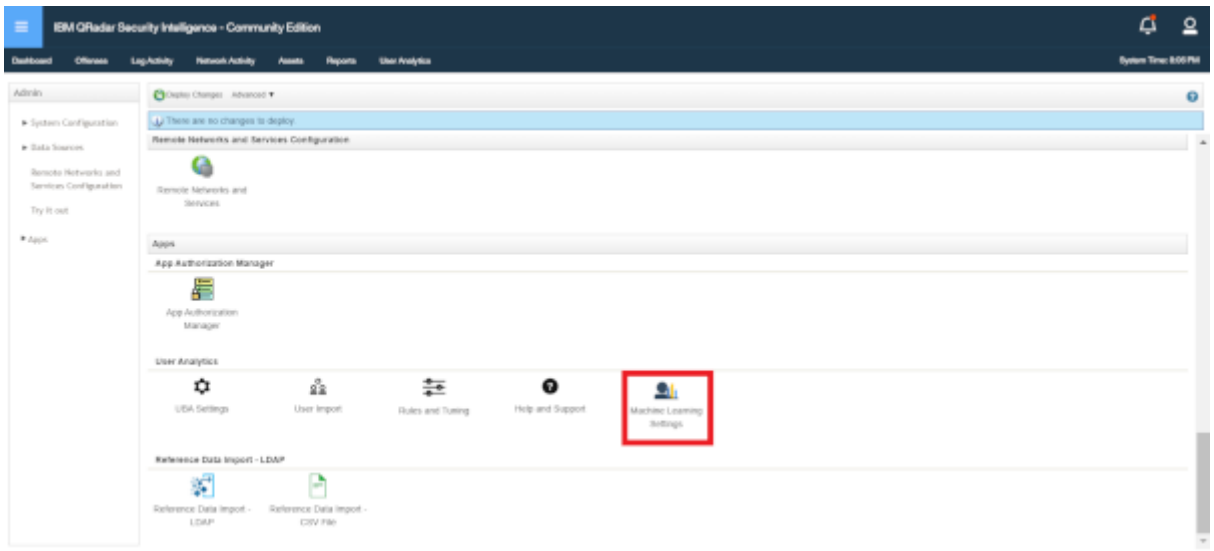

Figura C.1: Instalação ML App - Passo 1: Machine Learning Settings

Após esta ação é aberta uma janela, conforme figura C.2, onde são apresentadas as verificações feitas aos requisitos e, caso esteja tudo conforme, basta clicar no botão *Install ML App*.

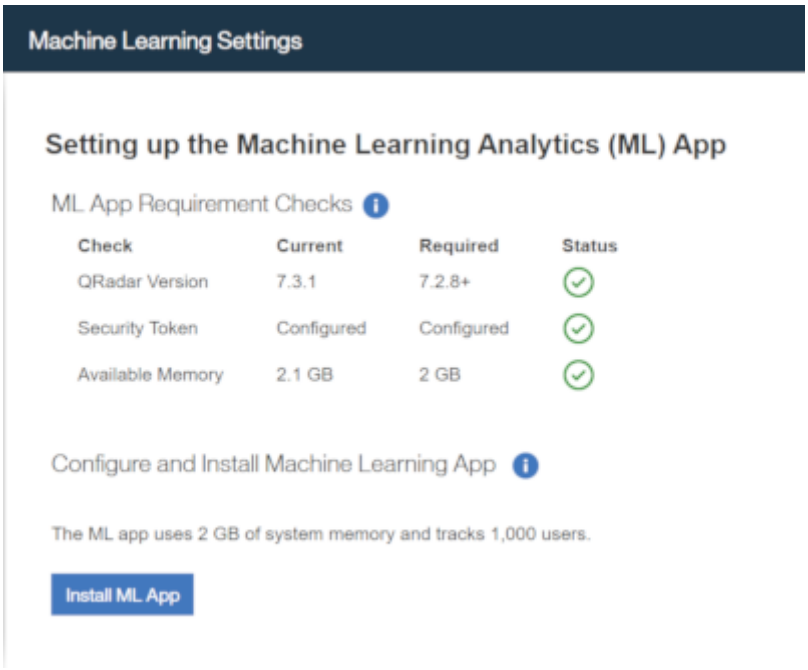

Figura C.2: Instalação ML App - Passo 2: Validação requisitos

De seguida, basta aceitar a instalação, a qual irá demorar alguns minutos.

Anexo D

Importação dos Dados dos Utilizadores da AD

Com o objetivo de obter informação de identificação contextual dos utilizadores, pode ser utilizada a *Reference Data Import - LDAP App*, a qual é instalada automaticamente durante a instalação da UBA App. Esta aplicação disponibiliza dois modos de importação de utilizadores: um através da ligação a um servidor da AD, e outro através de da utilização de um ficheiro CSV. O procedimento a seguir descrito foi executado fazendo uso deste segundo método:

1. No menu de navegação selecionar o separador *Admin*;
2. Na secção *Reference Data Import - LDAP*, clicar no ícone *Reference Data Import - CSV File*;

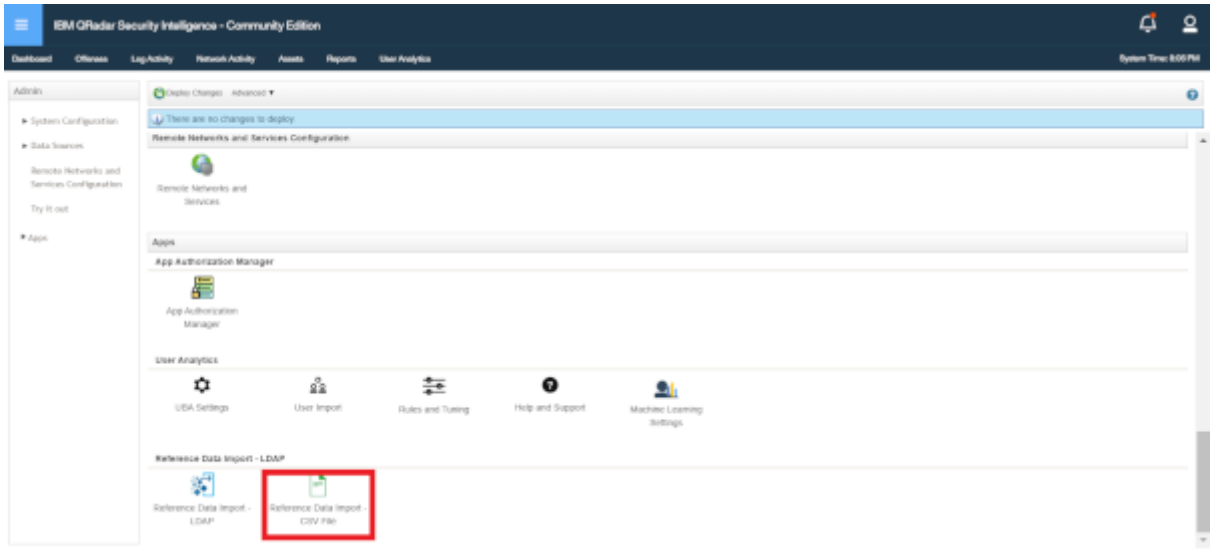

Figura D.1: Importação dos Dados dos Utilizadores da AD- Passo 1: Escolha método importação - ficheiro CSV

3. Na janela *Reference Data Import - CSV File*, apresentada na figura D.1, clicar em *Add Import*;
4. No ecrã da figura D.2, pesquisar a diretoria onde está guardado o ficheiro CSV com informação dos utilizadores. Este ficheiro deverá apresentar um tamanho inferior a 5 MB, conter um cabeçalho com o nome das colunas, e pelo menos uma coluna que contenha dados únicos;

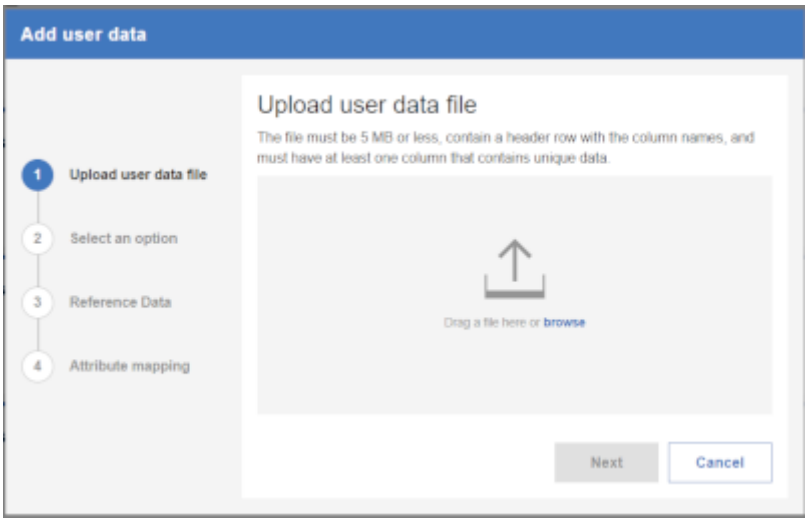

Figura D.2: Importação dos Dados dos Utilizadores da AD- Passo 2: Importação do ficheiro CSV

5. Clicar em seguinte e escolher se se pretende criar uma *Reference Table* ou unir os dados com uma já existente. Considerando que nesta importação se utilizou quatro ficheiros distintos (dizem respeito a 4 domínios diferentes), na importação do primeiro destes ficheiros selecionou-se a opção de criar uma *Reference Table*, enquanto que nas restantes três importações se uniu com esta mesma *Reference Table*;
6. Clicar novamente em seguinte, escolhendo de seguida os atributos desejados e a chave primária para essa mesma tabela, finalizando em *Import*.

Concluída esta primeira fase de importação, a fase que se segue vai permitir importar os utilizadores da *Reference Table*. Para tal, devem ser seguidos os seguintes passos:

1. No menu de navegação selecionar o separador *Admin*;
2. Na secção *User Analytics*, clicar no ícone *User Import*, conforme figura D.3;

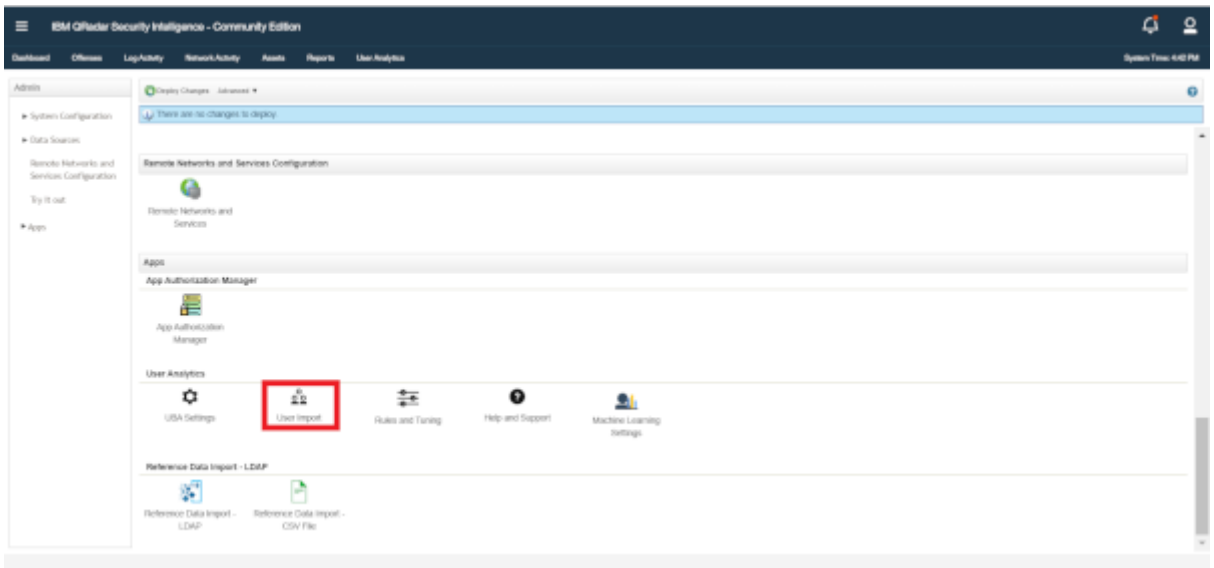

Figura D.3: Importação dos Dados dos Utilizadores da AD- Passo 3: User Import

3. Na janela da figura D.3, clicar em adicionar, escolhendo de seguida a opção *Reference Table*

apresentada na figura D.4;

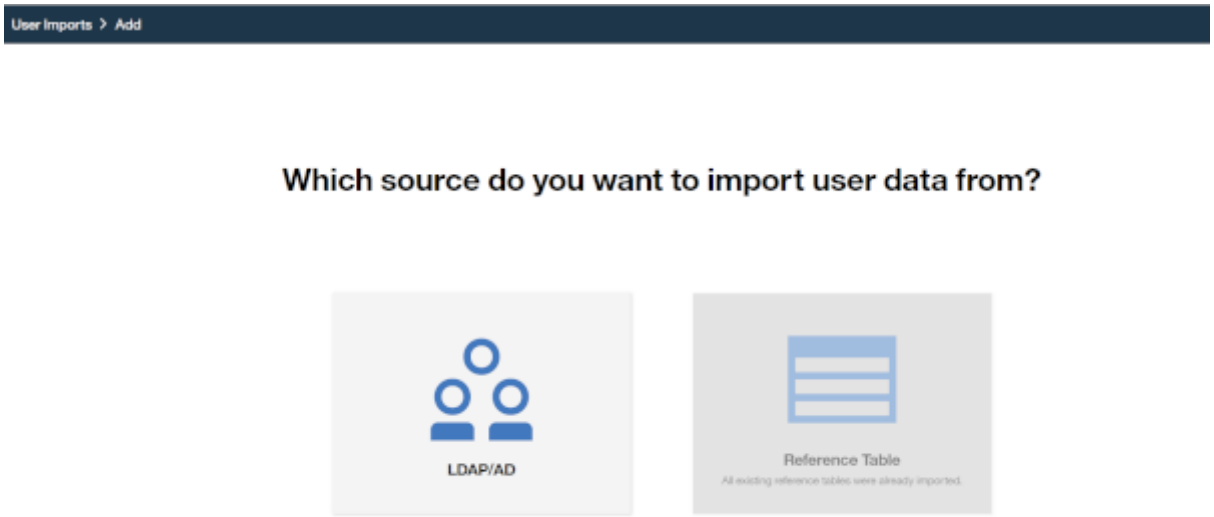

Figura D.4: Importação dos Dados dos Utilizadores da AD- Passo 4: Reference Table

4. No final, após a conclusão da importação, obtém-se a informação constante na figura D.5;

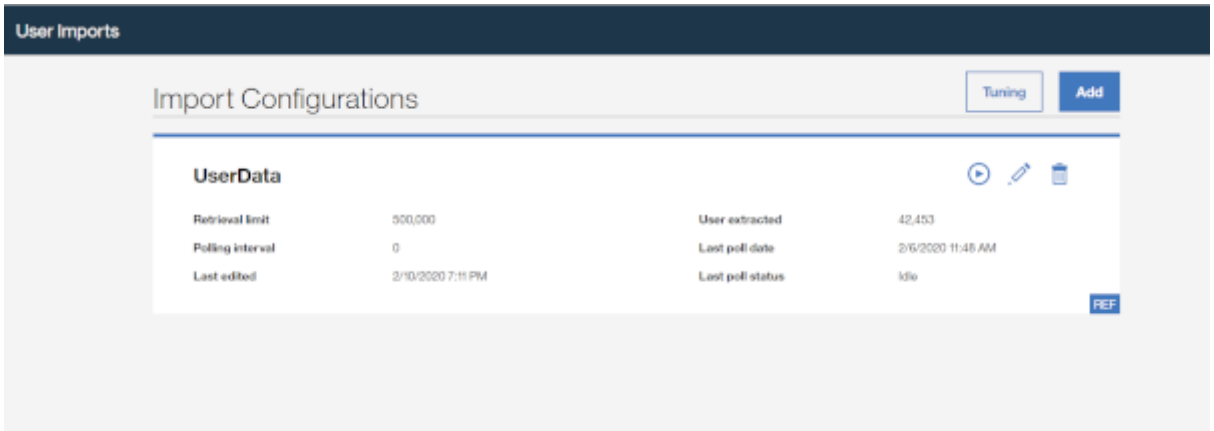

Figura D.5: Importação dos Dados dos Utilizadores da AD- Passo 5: Informação Resultante da Importação

5. De seguida, após conclusão das configurações de importação, afinar os atributos, recorrendo ao botão *Tuning*, presente nas figuras D.5;
6. Na secção *User Coalescing*, após clicar em editar, escolher os atributos *Logon Name* e *Email* para identificar e combinar atividades de diferentes nomes de utilizador, para cada utilizador. Na secção *Display Fields* efetuar a correspondência, conforme se ilustra na figura D.6;

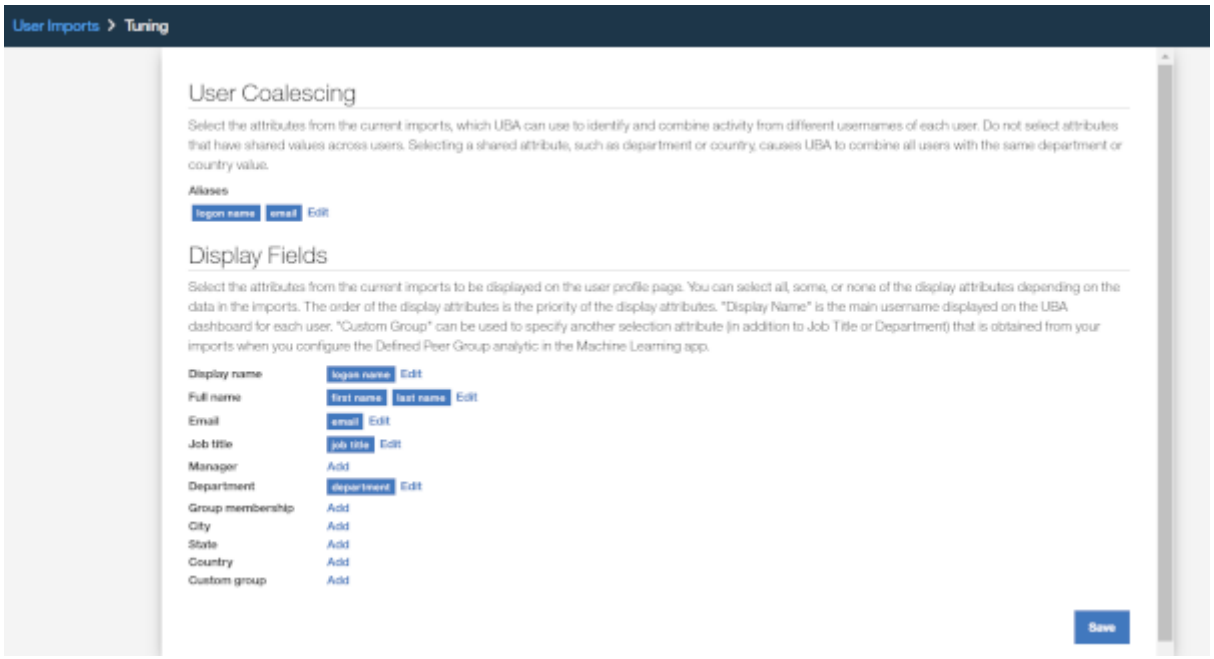

Figura D.6: Importação dos Dados dos Utilizadores da AD- Passo 6: User Coalescing

Anexo E

Visualização dos Dados dos Utilizadores da AD inseridos na *Reference Table*

Para visualizar o conteúdo da *reference table* e verificar quais os dados dos utilizadores da AD carregados e se os mesmos estão corretos, pode ser utilizada a interface de programação de aplicações (*Application Programming Interface* - API) disponibilizada pelo *QRadar*. Após invocação da mesma, o resultado final é obtido no formato JSON. Desta forma, detalha-se o procedimento a seguir:

1. No menu de navegação, escolher o separador *Interactive API for Developers*, o qual abrirá num novo separador;
2. Na versão de API mais recente, no presente caso 9.2 (as versões de API disponibilizadas dependem da versão de *QRadar* instalado), escolher *reference_data* > *tables* > *{name}*. Como parâmetros obrigatórios do pedido HTTP GET, colocar no campo *name* o nome da *reference table* configurado, neste caso *UserData*;

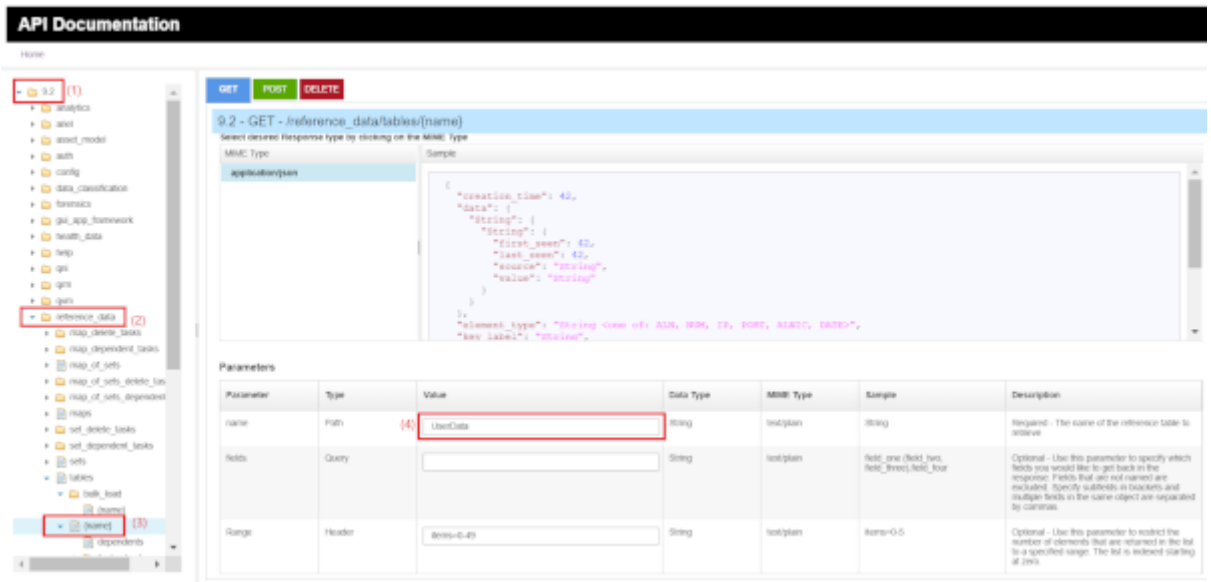

Figura E.1: Visualização Dados Utilizadores - Passo 1: Configuração do pedido HTTP

3. No final, depois de executado o pedido, em caso de sucesso, obtém-se os dados conforme a imagem seguinte.

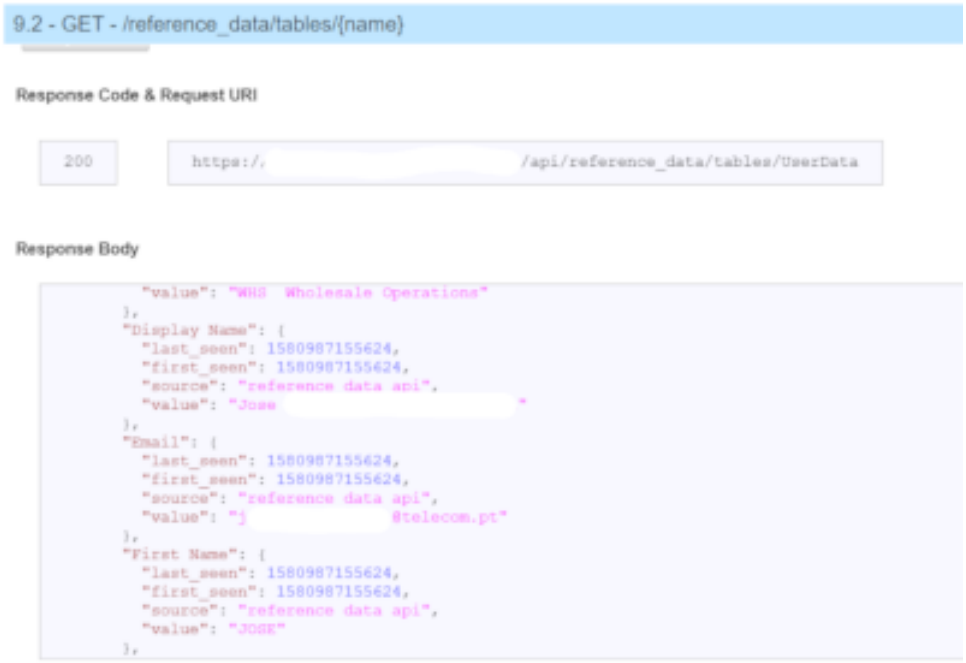

Figura E.2: Visualização Dados Utilizadores - Passo 2: Visualização da resposta ao pedido HTTP

Anexo F

Campos dos Eventos de Segurança das Estações *Windows* a Validar

Tabela F.1: Campos dos Eventos de Segurança das Estações Windows *a Validar*

Campo	**Origem**	
	Omissão	**Personalizado**
Event Name	X	
Low Level Category	X	
Event Description	X	
Username	X	
Start Time	X	
Storage Time	X	
Log Source Time	X	
Event ID		X
Logon Account Domain		X
Logon Account Name		X
Logon Type		X
Source Workstation		X
Status Code		X
Sub Status Code		X
Payload Information	X	
Event Hour		X

Anexo G

Filtro de Pesquisa AQL

```
(EventID='4624' AND ("Logon Type" = '2' OR "Logon Type" = '7' OR "Logon Type" = '10'
OR "Logon Type" = '11')) OR (EventID='4625' AND ("Sub Status Code" = '0xC0000070'
OR "Sub Status Code" = '0xC0000072' OR "Sub Status Code" = '0xC0000015B'
OR "Sub Status Code" = '0xC00000193') AND ("Source Workstation" <> username))
OR (EventID='4648') OR (EventID='4825')
```

Anexo H

Guião e Questionário para execução dos Testes

Guião para Execução do Teste I – Variante I

Tipo de Teste: Autenticação em estação de trabalho atribuída, no período habitual de trabalho.

Objetivo: Este teste procura validar o aparecimento, ou não, de falsos positivos.

Procedimento:

1. Realizar **uma** autenticação bem sucedida na estação de trabalho atribuída, durante o seu período habitual de trabalho.

   Nota: Ao autenticar-se na sua estação de trabalho, no início de um dia de trabalho, já está a realizar este teste, pelo que pode saltar imediatamente para o passo 2.

2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:

   (a) Indique o teste realizado no assunto do email (UBA_Teste1_Variante1).

   (b) Indique qual o seu utilizador.

   (c) Indique a estação de trabalho onde realizou a autenticação.

   Dica: Abrir linha de comandos > escrever *hostname*

   (d) A que horas realizou o teste?

Guião para Execução do Teste I – Variante II

Tipo de Teste: Autenticação em estação de trabalho atribuída, no período habitual de trabalho.

Objetivo: Este teste procura validar o aparecimento, ou não, de falsos positivos.

Procedimento:

1. Realizar **vinte ou mais** autenticações bem sucedidas na estação de trabalho atribuída, no intervalo de uma hora, durante o seu período habitual de trabalho.

   Nota: Ao autenticar-se na sua estação de trabalho, no início de um dia de trabalho, já está a realizar este uma autenticação, pelo que ficarão a faltar apenas 19 autenticações.

2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:

   (a) Indique o teste realizado no assunto do email (UBA_Teste1_Variante2).

   (b) Indique qual o seu utilizador.

   (c) Indique a estação de trabalho onde realizou a autenticação.

   Dica: Abrir linha de comandos > escrever *hostname*

   (d) A que horas realizou o teste?

Guião para Execução do Teste II – Variante I

Tipo de Teste: Autenticação em estação de trabalho não atribuída, no período habitual de trabalho.

Objetivo: Este teste destina-se a validar o correto funcionamento do caso de uso 1.

Procedimento:

1. Realizar **uma** autenticação bem sucedida numa estação de trabalho que **não** lhe esteja **atribuída**, durante o seu período habitual de trabalho.
2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:
   (a) Indique o teste realizado no assunto do email (UBA_Teste2_Variante1).
   (b) Indique qual o seu utilizador.
   (c) Indique a estação de trabalho onde realizou a autenticação.
   Dica: Abrir linha de comandos > escrever *hostname*
   (d) A que horas realizou o teste?

Guião para Execução do Teste II – Variante II

Tipo de Teste: Autenticação em estação de trabalho não atribuída, no período habitual de trabalho.

Objetivo: Este teste destina-se a validar o correto funcionamento do caso de uso 1.

Procedimento:

1. Realizar **cinco ou mais** autenticações bem sucedidas e/ou **cinco ou mais** tentativas de autenticação mal sucedidas (ex: falhar a palavra-passe) em estações de trabalho que **não** lhe estejam **atribuídas**, no intervalo de uma hora, durante o seu período habitual de trabalho. Cada autenticação e/ou tentativa de autenticação deve ser executada numa **estação de trabalho distinta**.
2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:
   (a) Indique o teste realizado no assunto do email (UBA_Teste2_Variante2).
   (b) Indique qual o seu utilizador.
   (c) Indique a estação de trabalho onde realizou a autenticação.
   Dica: Abrir linha de comandos > escrever *hostname*
   (d) A que horas realizou o teste?

Guião para Execução do Teste III – Variante I

Tipo de Teste: Autenticação em estação de trabalho atribuída, fora do período habitual de trabalho.

Objetivo: Este teste destina-se a validar o correto funcionamento do caso de uso 2.

Procedimento:

1. Realizar **uma** autenticação bem sucedida na estação de trabalho atribuída, **fora** do seu período habitual de trabalho.
2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:
   (a) Indique o teste realizado no assunto do email (UBA_Teste3_Variante1).
   (b) Indique qual o seu utilizador.
   (c) Indique a estação de trabalho onde realizou a autenticação.
   Dica: Abrir linha de comandos > escrever *hostname*
   (d) A que horas realizou o teste?

Guião para Execução do Teste III – Variante II

Tipo de Teste: Autenticação em estação de trabalho atribuída, fora do período habitual de trabalho.

Objetivo: Este teste destina-se a validar o correto funcionamento do caso de uso 2.

Procedimento:

1. Realizar **dez ou mais** autenticações bem sucedidas e/ou **dez ou mais** tentativas de autenticação mal sucedidas (ex: falhar a palavra-passe) na estação de trabalho **atribuída**, no intervalo de uma hora, **fora** do seu período habitual de trabalho.
2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:
   (a) Indique o teste realizado no assunto do email (UBA_Teste3_Variante2).
   (b) Indique qual o seu utilizador.
   (c) Indique a estação de trabalho onde realizou a autenticação.
   Dica: Abrir linha de comandos > escrever *hostname*
   (d) A que horas realizou o teste?

Guião para Execução do Teste IV – Variante I

Tipo de Teste: Autenticação em estação de trabalho não atribuída, fora do período habitual de trabalho.

Objetivo: Este teste destina-se a validar o correto funcionamento, em simultâneo, dos casos de uso 1 e 2.

Procedimento:

1. Realizar **uma** autenticação bem sucedida numa estação de trabalho que **não** lhe esteja **atribuída**, **fora** do seu período habitual de trabalho.
2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:
   (a) Indique o teste realizado no assunto do email (UBA_Teste4_Variante1).
   (b) Indique qual o seu utilizador.
   (c) Indique a estação de trabalho onde realizou a autenticação.
   Dica: Abrir linha de comandos > escrever *hostname*
   (d) A que horas realizou o teste?

Guião para Execução do Teste IV – Variante II

Tipo de Teste: Autenticação em estação de trabalho não atribuída, fora do período habitual de trabalho.

Objetivo: Este teste destina-se a validar o correto funcionamento, em simultâneo, dos casos de uso 1 e 2.

Procedimento:

1. Realizar **cinco ou mais** autenticações bem sucedidas e/ou **cinco ou mais** tentativas de autenticação mal sucedidas (ex: falhar a palavra-passe) em estações de trabalho que **não** lhe estejam **atribuídas**, no intervalo de uma hora, **fora** do seu período habitual de trabalho. Cada autenticação e/ou tentativa de autenticação deve ser executada numa **estação de trabalho distinta**.
2. Responder ao seguinte questionário, enviando um email para joao-g-filipe@telecom.pt:
   (a) Indique o teste realizado no assunto do email (UBA_Teste4_Variante2).
   (b) Indique qual o seu utilizador.
   (c) Indique a estação de trabalho onde realizou a autenticação.
   Dica: Abrir linha de comandos > escrever *hostname*
   (d) A que horas realizou o teste?

Anexo I

Criação de uma *Watchlist* e Seguimento de Utilizadores com ML

O procedimento aqui descrito detalha os passos realizados para a criação de uma *Watchlist* a partir de um *Reference Set*, bem como o método que permite forçar o seguimento, de um dado utilizador, através dos modelos analíticos de *Machine Learning*.

1. Criação de um *Reference Set* com os utilizadores pretendidos

   (a) No menu de navegação selecionar o separador *Admin*;

   (b) Clicar no ícone *Reference Set Management (System Configuration > Reference Set Management)*;

   (c) Após entrar nesta nova página, clicar no botão para adicionar um novo *reference set*. Preencher o campo com o nome desejado, bem como selecionado o tipo de dados (alfanumérico, ignorando o facto de os caracteres serem maiúsculos ou minúsculos), conforme ilustrado na figura I.1;

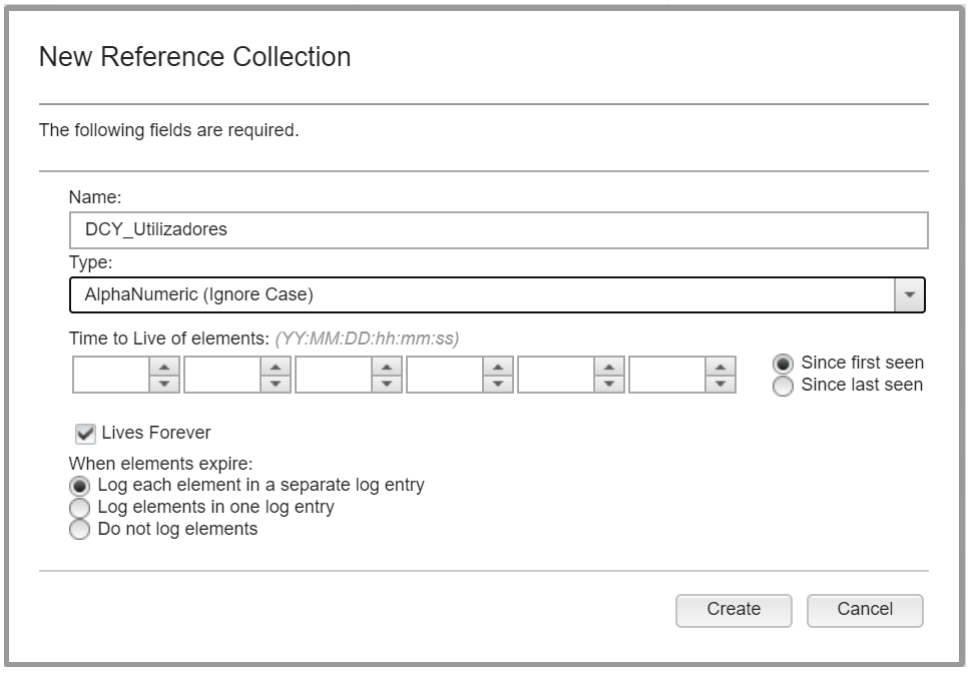

Figura I.1: Criação do Reference Set

(d) Após a criação do novo *reference set*, é preciso selecioná-lo novamente e visualizar os seus conteúdos, carregando depois no botão importar, o qual permite carregar para este *reference set*, um conjunto de utilizadores existentes num ficheiro CSV;

(e) Após a importação terminar, o trabalho de criação e configuração do *reference set* fica concluído, sendo possível visualizar o resumo dos seus dados, conforme figura I.2.

Name	Type	Number of Elements
DCY_Utilizadores	AlphaNumeric	44

Figura I.2: Estado final da criação do Reference Set

2. Criação de uma *Watchlist*

(a) A partir do UBA *Dashboard* ou da página de detalhes do utilizador, clicar no ícone *watchlist*, escolhendo de seguida no menu a opção *Create new watchlist*;

(b) No primeiro separador da nova janela, *General Settings*, é possível introduzir o nome da nova *watchlist* (DCY Watchlist), definir o fator de escalonamento do risco (escolhe-se o valor um, de modo a não apresentar qualquer influência) e a prioridade de seguimento com ML (escolhe-se o valor de alta), como se observa na figura I.3;

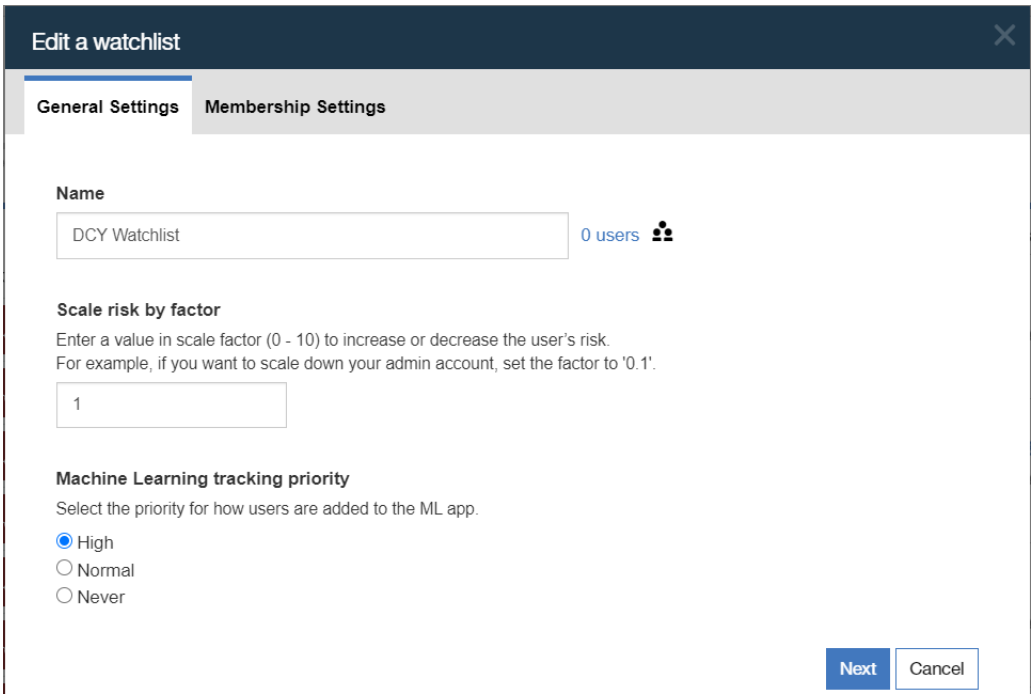

Figura I.3: Criação da Watchlist - Passo 1

(c) Clicando no botão avançar, passa-se para o segundo separador, *Membership Settings*. Neste separador é possível importar o *reference set* anteriormente criado na figura I.1, bem como definir o intervalo de atualização da referida *watchlist*. A figura I.4 demonstra as configurações introduzidas.

3. Seguimento de um Utilizador com ML

(a) A partir da página de detalhes do utilizador, clicar no botão *Advanced Actions*, esco-

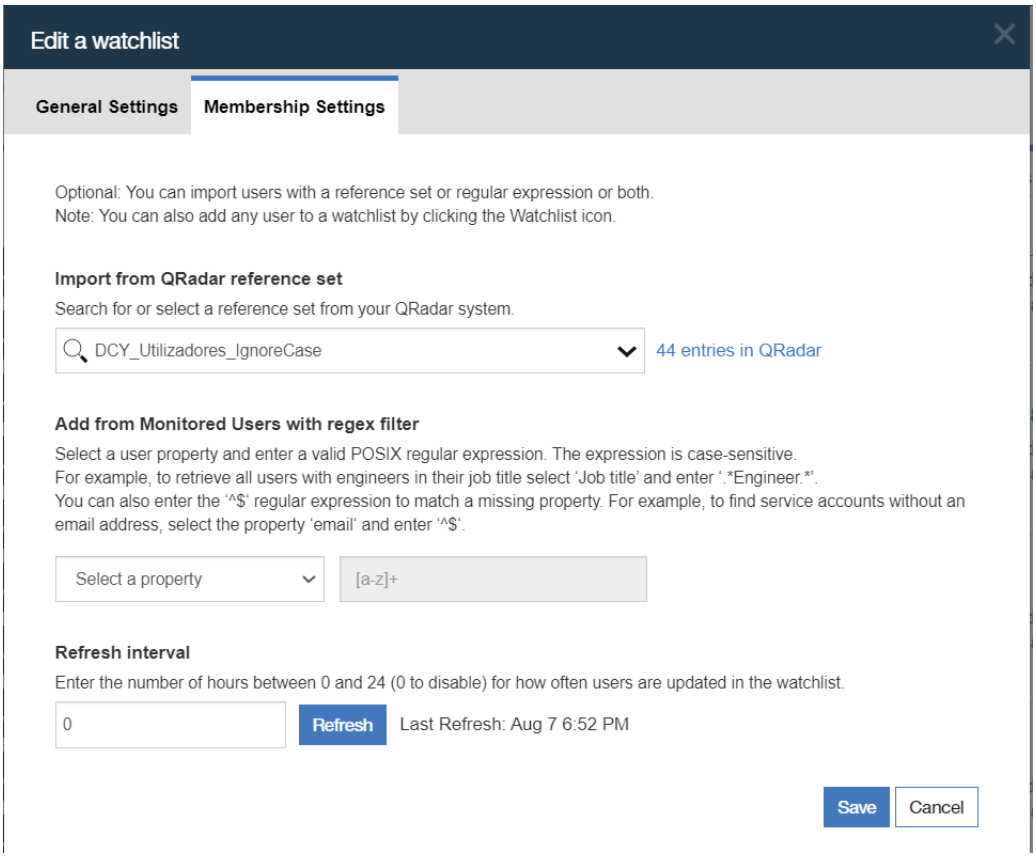

Figura I.4: Criação da Watchlist - Passo 2

lhendo depois no menu a opção *Always track with Machine Learning*;

(b) Após esta opção ser selecionada, o seguimento fica ativado, podendo ser confirmado novamente no mesmo menu, conforme figura I.5.

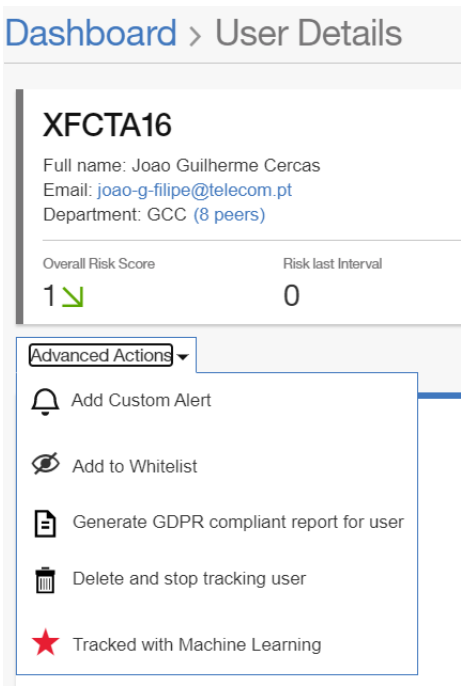

Figura I.5: Verificação do seguimento com ML de um utilizador

Anexo J

Scripts para Geração de Eventos de Autenticação

Script para Geração de Eventos do Tipo 4624

```
ï»¿#!/bin/bash

#Definir variaveis globais
USERNAME="XFCTA16"
DOMAIN="PTPORTUGAL"
SOURCE_WORKSTATION="ET81000259"
PATH_LOG_FILENAME="/home/xfcta16/log_4624.txt"
NR_EVENTS_PER_HOUR=$(( RANDOM % 10 )  + 1 ))
echo "Numero Eventos por Hora: "$NR_EVENTS_PER_HOUR
echo "Data/Hora Inicio: "$(date)


#Funcao que permite executa o envio de eventos para o QRadar
function send_syslog_event {
    echo "Numero Eventos a Enviar: "$1
    echo "Tempo para Enviar (minutos): "$2

    sleep $2m
    rm -f $PATH_LOG_FILENAME
    for (( c=1; c<=$1; c++ ))
        do
            write_syslog_event_4624_to_file
        done
    inject_syslog_events_to_qradar
}

#Funcao que escreve um evento 4624 num ficheiro, no grupo-data-hora definido
function write_syslog_event_4624_to_file {
    LANG=C
    local date_short=$(date '+%b %d %T')
    local date_full=$(date '+%a %b %d %T %Y')
    local EVENT_4624="<1>${date_short} ${SOURCE_WORKSTATION}.ptportugal.corppt.
        com MSWinEventLog       1       Security        186387  ${date_full}
            4624    Microsoft-Windows-Security-Auditing     N/A     N/A
        Success Audit   ${SOURCE_WORKSTATION}.ptportugal.corppt.com     Logon
                An account was successfully logged on.    Subject:   Security ID
        :  S-1-5-18   Account Name:  ${SOURCE_WORKSTATION}$   Account Domain:  $
```

```
        {DOMAIN}   Logon ID:  0x3E7    Logon Information:   Logon Type:  7
        Restricted Admin Mode: -   Virtual Account:  No   Elevated Token:  Yes
           Impersonation Level:  Impersonation    New Logon:   Security ID:  S
        -1-5-21-507921405-1336601894-725345543-1961   Account Name:  ${USERNAME}
           Account Domain:  ${DOMAIN}   Logon ID:  0x38435C1   Linked Logon ID:
         0x3843665   Network Account Name: -   Network Account Domain: -   Logon
         GUID:  {00000000-0000-0000-0000-000000000000}    Process Information:
          Process ID:  0x770   Process Name:  C:\Windows\System32\svchost.exe
           Network Information:   Workstation Name: ${SOURCE_WORKSTATION}
        Source Network Address: 127.0.0.1   Source Port:  0    Detailed Aut"
    echo $EVENT_4624 >> $PATH_LOG_FILENAME
}

#Funcao que permite injetar eventos no QRadar a partir de ficheiro
function inject_syslog_events_to_qradar {
    /opt/qradar/bin/logrun.pl -f $PATH_LOG_FILENAME 1
}

#Invocacao das funcoes anteriores
send_syslog_event $NR_EVENTS_PER_HOUR $(( RANDOM % 55 )  + 1 ))
```

Script para Geração de Eventos do Tipo 4625

```
ï»¿#!/bin/bash

#Definir variaveis globais
USERNAME="XFCTA16"
DOMAIN="PTPORTUGAL"
SOURCE_WORKSTATION="ET81000259"
PATH_LOG_FILENAME="/home/xfcta16/log_4625.txt"
NR_EVENTS_PER_DAY=$(( RANDOM % 3 ))
SEND_WITHIN_HOUR=$(( RANDOM % 10 ))
SEND_MINUTE=$(( RANDOM % 59 )  + 1 ))
echo "Numero Eventos 4625 Diarios: "$NR_EVENTS_PER_DAY
echo "Hora Envio: "$SEND_WITHIN_HOUR
echo "Minuto Envio: "$SEND_MINUTE
echo "Data/Hora Inicio: "$(date)


#Funcao que executa a criacao e envio de eventos para o QRadar
function send_syslog_event {
    for (( c=1; c<=$1; c++ ))
        do
            write_syslog_event_4625_to_file
        done
    write_syslog_event_4624_to_file
    inject_syslog_events_to_qradar
}

#Funcao que escreve um evento 4625 num ficheiro, no grupo-data-hora definido
function write_syslog_event_4625_to_file {
    LANG=C
    local date_short=$(date '+%b %d %T')
    local date_full=$(date '+%a %b %d %T %Y')
```

```
32      local EVENT_4625="<11>${date_short} ${SOURCE_WORKSTATION}.ptportugal.corppt.
            com MSWinEventLog       3       Security        165551  ${date_full}
            4625    Microsoft-Windows-Security-Auditing     N/A     N/A     Failure
            Audit   ${SOURCE_WORKSTATION}.ptportugal.corppt.com     Logon
            An account failed to log on.    Subject:   Security ID:  S-1-5-18
            Account Name:  ${SOURCE_WORKSTATION}$   Account Domain:  ${DOMAIN}
            Logon ID:  0x3E7    Logon Type:   7    Account For Which Logon Failed:
              Security ID:  S-1-0-0   Account Name:  ${USERNAME}   Account Domain:
            ${DOMAIN}    Failure Information:   Failure Reason:  Unknown user name
            or bad password.   Status:   0xC000006D   Sub Status:  0xC000006A
            Process Information:   Caller Process ID: 0x740   Caller Process Name: C
            :\Windows\System32\svchost.exe    Network Information:   Workstation
            Name: ${SOURCE_WORKSTATION}   Source Network Address: 127.0.0.1   Source
             Port:  0    Detailed Authentication Information:   Logon Process:
            User32    Authentication Package: Negotiate   Transited Services: -
            Package Name (NTLM only): -   Key Length:  0    This event is ge"
33      echo $EVENT_4625 >> $PATH_LOG_FILENAME
34  }
35
36  #Funcao que escreve um evento 4624 num ficheiro, no grupo-data-hora definido
37  function write_syslog_event_4624_to_file {
38      LANG=C
39      local date_short=$(date '+%b %d %T')
40      local date_full=$(date '+%a %b %d %T %Y')
41      local EVENT_4624="<1>${date_short} ${SOURCE_WORKSTATION}.ptportugal.corppt.
            com MSWinEventLog       1       Security        186387  ${date_full}
                4624    Microsoft-Windows-Security-Auditing     N/A     N/A
            Success Audit   ${SOURCE_WORKSTATION}.ptportugal.corppt.com     Logon
                    An account was successfully logged on.    Subject:   Security ID
            :  S-1-5-18   Account Name:  ${SOURCE_WORKSTATION}$   Account Domain:  $
            {DOMAIN}   Logon ID:  0x3E7    Logon Information:   Logon Type:  7
            Restricted Admin Mode: -   Virtual Account:  No   Elevated Token:  Yes
               Impersonation Level:  Impersonation    New Logon:   Security ID:  S
            -1-5-21-507921405-1336601894-725345543-1961   Account Name:  ${USERNAME}
               Account Domain:  ${DOMAIN}   Logon ID:  0x38435C1   Linked Logon ID:
             0x3843665   Network Account Name: -   Network Account Domain: -   Logon
             GUID:  {00000000-0000-0000-0000-000000000000}    Process Information:
              Process ID:  0x770   Process Name:  C:\Windows\System32\svchost.exe
               Network Information:   Workstation Name: ${SOURCE_WORKSTATION}
            Source Network Address: 127.0.0.1   Source Port:  0    Detailed Aut"
42      echo $EVENT_4624 >> $PATH_LOG_FILENAME
43  }
44
45  #Funcao que permite injetar eventos no QRadar a partir de um ficheiro
46  function inject_syslog_events_to_qradar {
47      /opt/qradar/bin/logrun.pl -f $PATH_LOG_FILENAME 1
48      unlink $PATH_LOG_FILENAME
49  }
50
51  #Invocacao das funcoes anteriores
52  if [ $NR_EVENTS_PER_DAY -ne 0 ] ; then
53      sleep $(((SEND_WITHIN_HOUR*60) +SEND_MINUTE))
54      send_syslog_event $NR_EVENTS_PER_DAY
55  fi
```

Script para Geração de Eventos Anómalos

```
#!/bin/bash

#Definir variaveis globais
USERNAME="XFCTA16"
DOMAIN="PTPORTUGAL"
SOURCE_WORKSTATION="TD00826112"
PATH_LOG_FILENAME="/home/xfcta16/log_anomalos.txt"


#Funcao que executa a criacao e envio de eventos para o QRadar
function send_syslog_event {
    rm -f $PATH_LOG_FILENAME
    for (( c=1; c<=$1; c++ ))
        do
            if [ $2 -eq 4624 ] ; then
                write_syslog_event_4624_to_file
            else
                write_syslog_event_4625_to_file
            fi
        done
    inject_syslog_events_to_qradar
}

#Funcao que escreve um evento 4625 num ficheiro, no grupo-data-hora definido
function write_syslog_event_4625_to_file {
    LANG=C
    local date_short=$(date '+%b %d %T')
    local date_full=$(date '+%a %b %d %T %Y')
    local EVENT_4625="<11>${date_short} ${SOURCE_WORKSTATION}.ptportugal.corppt.
        com MSWinEventLog       3       Security        165551  ${date_full}
        4625    Microsoft-Windows-Security-Auditing     N/A     N/A     Failure
        Audit   ${SOURCE_WORKSTATION}.ptportugal.corppt.com     Logon
        An account failed to log on.    Subject:   Security ID:  S-1-5-18
        Account Name:  ${SOURCE_WORKSTATION}$   Account Domain:  ${DOMAIN}
        Logon ID:  0x3E7    Logon Type:   7    Account For Which Logon Failed:
          Security ID:  S-1-0-0   Account Name:  ${USERNAME}   Account Domain:
        ${DOMAIN}    Failure Information:   Failure Reason:  Unknown user name
        or bad password.   Status:   0xC000006D   Sub Status:  0xC000006A
        Process Information:   Caller Process ID: 0x740   Caller Process Name: C
        :\Windows\System32\svchost.exe    Network Information:   Workstation
        Name: ${SOURCE_WORKSTATION}   Source Network Address: 127.0.0.1   Source
         Port:  0    Detailed Authentication Information:   Logon Process:
        User32    Authentication Package: Negotiate   Transited Services: -
        Package Name (NTLM only): -   Key Length:  0    This event is ge"
    echo $EVENT_4625 >> $PATH_LOG_FILENAME
}

#Funcao que escreve um evento 4624 num ficheiro, no grupo-data-hora definido
function write_syslog_event_4624_to_file {
    LANG=C
    local date_short=$(date '+%b %d %T')
    local date_full=$(date '+%a %b %d %T %Y')
    local EVENT_4624="<1>${date_short} ${SOURCE_WORKSTATION}.ptportugal.corppt.
        com MSWinEventLog       1       Security        186387  ${date_full}
            4624    Microsoft-Windows-Security-Auditing     N/A     N/A
        Success Audit   ${SOURCE_WORKSTATION}.ptportugal.corppt.com     Logon
```

```
                An account was successfully logged on.    Subject:   Security ID
        :  S-1-5-18   Account Name:  ${SOURCE_WORKSTATION}$   Account Domain:  $
        {DOMAIN}   Logon ID:  0x3E7    Logon Information:   Logon Type:  7
        Restricted Admin Mode: -   Virtual Account:  No   Elevated Token:  Yes
           Impersonation Level:  Impersonation    New Logon:   Security ID:  S
        -1-5-21-507921405-1336601894-725345543-1961   Account Name:  ${USERNAME}
           Account Domain:  ${DOMAIN}   Logon ID:  0x38435C1   Linked Logon ID:
         0x3843665   Network Account Name: -   Network Account Domain: -   Logon
         GUID:  {00000000-0000-0000-0000-000000000000}    Process Information:
          Process ID:  0x770   Process Name:  C:\Windows\System32\svchost.exe
           Network Information:   Workstation Name: ${SOURCE_WORKSTATION}
        Source Network Address: 127.0.0.1   Source Port:  0    Detailed Aut"
    echo $EVENT_4624 >> $PATH_LOG_FILENAME
}

#Funcao que permite injetar eventos no QRadar a partir de um ficheiro
function inject_syslog_events_to_qradar {
    /opt/qradar/bin/logrun.pl -f $PATH_LOG_FILENAME 1
    echo "Terminado envio dos eventos!"
}

#Executa a leitura dos parametros de entrada e invoca as funcoes anteriores
echo -e "Introduza o tipo evento que pretende enviar: \c "
read event_type
echo -e "Introduza o numero de eventos que pretende enviar: \c "
read nr_events
send_syslog_event $nr_events $event_type
```

Configuração do Escalonador de Tarefas do Sistema Operativo

```
0 8-19 * * mon-fri root bash /home/xfcta16/geracao_eventos_4624.sh
0 8 * * mon-fri root bash /home/xfcta16/geracao_eventos_4625.sh
```